

OM of: US-09-528-682-2 to: GenEmbl:\* out\_format : pfs  
Date: Jun 18, 2002 7:27 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+ p2n\_model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09528682/runat\_18062002\_082443\_7808/app\_query.fasta\_1.689  
-DB=genEmbl -QFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=1000 -DOALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682\_CGNI\_1\_8225 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-2  
Query length: 3  
Database: GenEmbl:\*  
Database sequences: 1797656  
Database length: 187333701  
Search time (sec): 3690.420000

score\_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation	...
gb_pat:AR096851	+	16.00	105.33	1.7e+03	10	AR096851	Sequence 2 from patent
gb_pat:AR096852	+	16.00	105.33	1.7e+03	10	AR096852	Sequence 3 from patent
gb_pat:AR096853	-	16.00	105.33	1.7e+03	10	AR096853	Sequence 110 from patent
gb_pat:AR096854	-	16.00	105.33	1.7e+03	10	AR096854	Sequence 235 from patent
gb_pat:AR096855	+	16.00	104.69	1.8e+03	11	AR096855	Sequence 54 from patent
gb_pat:AR096856	+	16.00	104.69	1.8e+03	11	AR096856	Sequence 3 from patent
gb_pat:AR096857	-	16.00	104.11	2.0e+03	12	AR096857	Sequence 13 from patent
gb_pat:AR096858	+	16.00	104.11	2.0e+03	12	AR096858	Sequence 9 from patent
gb_pat:AR096859	+	16.00	104.11	2.0e+03	12	AR096859	Sequence 9 from patent
gb_pat:AR096860	+	16.00	104.11	2.0e+03	12	AR096860	Sequence 8 from patent
gb_pat:AR096861	-	16.00	103.07	2.2e+03	14	AR096861	Sequence 211 from patent
gb_pat:AR096862	-	16.00	103.07	2.2e+03	14	AR096862	Sequence 83 from patent
gb_pat:AR096863	-	16.00	103.07	2.2e+03	14	AR096863	Sequence 3 from patent
gb_pat:AR096864	+	16.00	103.07	2.2e+03	14	AR096864	Sequence 19 from patent
gb_pat:AR096865	-	16.00	103.07	2.2e+03	14	AR096865	Sequence 7 from patent
gb_pat:AR096866	-	16.00	103.07	2.2e+03	14	AR096866	Sequence 66 from patent
gb_pat:AR096867	+	16.00	103.07	2.2e+03	14	AR096867	Sequence 211 from patent
gb_pat:AR096868	+	16.00	103.07	2.2e+03	14	AR096868	Sequence 213 from patent
gb_pat:AR096869	+	16.00	103.07	2.2e+03	14	AR096869	Sequence 255 from patent
gb_pat:AR096870	+	16.00	103.07	2.2e+03	14	AR096870	Sequence 258 from patent
gb_pat:AR096871	+	16.00	103.07	2.2e+03	14	AR096871	Sequence 5 from patent
gb_pat:AR096872	+	16.00	103.07	2.2e+03	14	AR096872	Sequence 6 from patent
gb_pat:AR096873	+	16.00	103.07	2.2e+03	14	AR096873	Sequence 5 from patent
gb_pat:AR096874	+	16.00	103.07	2.2e+03	14	AR096874	Sequence 11 from patent
gb_pat:AR096875	+	16.00	103.07	2.2e+03	14	AR096875	Sequence 164 from patent
gb_pat:AR096876	+	16.00	103.07	2.2e+03	14	AR096876	Sequence 8 from patent
gb_pat:AR096877	+	16.00	103.07	2.2e+03	14	AR096877	Sequence 5 from patent
gb_pat:AR096878	+	16.00	103.07	2.2e+03	14	AR096878	Sequence 6 from patent
gb_pat:AR096879	-	16.00	102.61	2.4e+03	15	AR096879	Sequence 2 from patent
gb_pat:AR096880	+	16.00	102.61	2.4e+03	15	AR096880	Sequence 254 from patent
gb_pat:AR096881	+	16.00	102.61	2.4e+03	15	AR096881	Sequence 256 from patent
gb_pat:AR096882	+	16.00	102.61	2.4e+03	15	AR096882	Sequence 355 from patent
gb_pat:AR096883	-	16.00	102.61	2.4e+03	15	AR096883	Sequence 355 from patent



gb_pat:AX284076	16	16.00	102.18	2.5e+03	16	AX284076	Sequence 41	from Patent	gb_pat:154062	17	16.00	101.77	2.6e+03	17	I54062	Sequence 1803	from pat
gb_pat:AX300970	16	16.00	102.18	2.5e+03	16	AX300970	Sequence 41	from Patent	gb_pat:154124	17	16.00	101.77	2.6e+03	17	I54124	Sequence 1865	from pat
gb_pat:141042	16	16.00	102.18	2.5e+03	16	141042	Sequence 11	from patent	gb_pat:194434	17	16.00	101.77	2.6e+03	17	I94434	Sequence 597	from pat
gb_pat:141043	16	16.00	102.18	2.5e+03	16	141043	Sequence 12	from patent	em_pat:E09908	17	16.00	101.77	2.6e+03	17	E09908	Probe 17T-2	for detect
gb_pat:147692	16	16.00	102.18	2.5e+03	16	147692	Sequence 4	from patent	gb_pat:AX31539	18	16.00	101.39	2.8e+03	18	AX31539	Synthetic H.pylori det	
gb_pat:AR016252	17	16.00	101.77	2.6e+03	17	AR016252	Sequence 6	from patent	gb_pat:AR002996	18	16.00	101.39	2.8e+03	18	AR002996	Sequence 35	from pat
gb_pat:AR039563	17	16.00	101.77	2.6e+03	17	AR039563	Sequence 411	from patent	gb_pat:AR011218	18	16.00	101.39	2.8e+03	18	AR011218	Sequence 86	from pat
gb_pat:AR039565	17	16.00	101.77	2.6e+03	17	AR039565	Sequence 413	from patent	gb_pat:AR053147	18	16.00	101.39	2.8e+03	18	AR053147	Sequence 53	from pat
gb_pat:AR039567	17	16.00	101.77	2.6e+03	17	AR039567	Sequence 415	from patent	gb_pat:AR144877	18	16.00	101.39	2.8e+03	18	AR144877	Sequence 122	from pa
gb_pat:AR039571	17	16.00	101.77	2.6e+03	17	AR039571	Sequence 417	from patent	gb_pat:AX048677	18	16.00	101.39	2.8e+03	18	AX048677	Sequence 122	from Pat
gb_pat:AR039573	17	16.00	101.77	2.6e+03	17	AR039573	Sequence 419	from patent	gb_pat:AX076467	18	16.00	101.39	2.8e+03	18	AX076467	Sequence 12	from Pat
gb_pat:AR039577	17	16.00	101.77	2.6e+03	17	AR039577	Sequence 421	from patent	gb_pat:AX078832	18	16.00	101.39	2.8e+03	18	AX078832	Sequence 6	from Pat
gb_pat:AR045997	17	16.00	101.77	2.6e+03	17	AR045997	Sequence 790	from patent	gb_pat:AX078833	18	16.00	101.39	2.8e+03	18	AX078833	Sequence 7	from Pat
gb_pat:AR046273	17	16.00	101.77	2.6e+03	17	AR046273	Sequence 1066	from patent	gb_pat:AX078834	18	16.00	101.39	2.8e+03	18	AX078834	Sequence 8	from Pat
gb_pat:AR046275	17	16.00	101.77	2.6e+03	17	AR046275	Sequence 1068	from patent	gb_pat:AX078843	18	16.00	101.39	2.8e+03	18	AX078843	Sequence 17	from Pat
gb_pat:AR046277	17	16.00	101.77	2.6e+03	17	AR046277	Sequence 1072	from patent	gb_pat:AX078849	18	16.00	101.39	2.8e+03	18	AX078849	Sequence 23	from Pat
gb_pat:AR046279	17	16.00	101.77	2.6e+03	17	AR046279	Sequence 1074	from patent	gb_pat:AX078860	18	16.00	101.39	2.8e+03	18	AX078860	Sequence 34	from Pat
gb_pat:AR046281	17	16.00	101.77	2.6e+03	17	AR046281	Sequence 1074	from patent	gb_pat:AX135661	18	16.00	101.39	2.8e+03	18	AX135661	Sequence 2	from Pat
gb_pat:AR046283	17	16.00	101.77	2.6e+03	17	AR046283	Sequence 1076	from patent	gb_pat:AX135661	18	16.00	101.39	2.8e+03	18	AX135661	Sequence 2	from Pat
gb_pat:AR046285	17	16.00	101.77	2.6e+03	17	AR046285	Sequence 1078	from patent	gb_pat:AX191970	18	16.00	101.39	2.8e+03	18	AX191970	Sequence 122	from Pa
gb_pat:AR046998	17	16.00	101.77	2.6e+03	17	AR046998	Sequence 1791	from patent	gb_pat:BD009430	18	16.00	101.39	2.8e+03	18	BD009430	Probes, methods and	
gb_pat:AR047000	17	16.00	101.77	2.6e+03	17	AR047000	Sequence 1791	from patent	gb_pat:E08099	18	16.00	101.39	2.8e+03	18	E08099	Synthetic RNA fragment	
gb_pat:AR047002	17	16.00	101.77	2.6e+03	17	AR047002	Sequence 1795	from patent	gb_pat:E32450	18	16.00	101.39	2.8e+03	18	E32450	Mammal-derived tissue	
gb_pat:AR047004	17	16.00	101.77	2.6e+03	17	AR047004	Sequence 1795	from patent	gb_pat:E32451	18	16.00	101.39	2.8e+03	18	E32451	Mammal-derived tissue	
gb_pat:AR047006	17	16.00	101.77	2.6e+03	17	AR047006	Sequence 1797	from patent	gb_pat:E32452	18	16.00	101.39	2.8e+03	18	E32452	Mammal-derived tissue	
gb_pat:AR047008	17	16.00	101.77	2.6e+03	17	AR047008	Sequence 1801	from patent	gb_pat:E32453	18	16.00	101.39	2.8e+03	18	E32453	Mammal-derived tissue	
gb_pat:AR047010	17	16.00	101.77	2.6e+03	17	AR047010	Sequence 1803	from patent	gb_pat:E32454	18	16.00	101.39	2.8e+03	18	E32454	Mammal-derived tissue	
gb_pat:AR047072	17	16.00	101.77	2.6e+03	17	AR047072	Sequence 1865	from patent	gb_pat:E32455	18	16.00	101.39	2.8e+03	18	E32455	Mammal-derived tissue	
gb_pat:AR04983	17	16.00	101.77	2.6e+03	17	AR04983	Sequence 21	from patent	gb_pat:E32456	18	16.00	101.39	2.8e+03	18	E32456	Mammal-derived tissue	
gb_pat:AR101754	17	16.00	101.77	2.6e+03	17	AR101754	Sequence 21	from patent	gb_pat:E32457	18	16.00	101.39	2.8e+03	18	E32457	Mammal-derived tissue	
gb_pat:AR166243	17	16.00	101.77	2.6e+03	17	AR166243	Sequence 16	from patent	gb_pat:E32458	18	16.00	101.39	2.8e+03	18	E32458	Mammal-derived tissue	
gb_pat:AX216730	17	16.00	101.77	2.6e+03	17	AX216730	Sequence 16	from patent	gb_pat:E32459	18	16.00	101.39	2.8e+03	18	E32459	Mammal-derived tissue	
gb_pat:AX217526	17	16.00	101.77	2.6e+03	17	AX217526	Sequence 2172	from patent	gb_pat:E32460	18	16.00	101.39	2.8e+03	18	E32460	Mammal-derived tissue	
gb_pat:AX217527	17	16.00	101.77	2.6e+03	17	AX217527	Sequence 2968	from patent	gb_pat:E32461	18	16.00	101.39	2.8e+03	18	E32461	Mammal-derived tissue	
gb_pat:AX217528	17	16.00	101.77	2.6e+03	17	AX217528	Sequence 2969	from patent	gb_pat:I04015	18	16.00	101.39	2.8e+03	18	I04015	Sequence 11	from Patent
gb_pat:AX217529	17	16.00	101.77	2.6e+03	17	AX217529	Sequence 2970	from patent	gb_pat:I17856	18	16.00	101.39	2.8e+03	18	I17856	Sequence 86	from patent
gb_pat:AX217530	17	16.00	101.77	2.6e+03	17	AX217530	Sequence 2971	from patent	gb_pat:I73187	18	16.00	101.39	2.8e+03	18	I73187	Sequence 1	from patent
gb_pat:AX217531	17	16.00	101.77	2.6e+03	17	AX217531	Sequence 2972	from patent	gb_pat:A64247	19	16.00	101.03	2.9e+03	19	A64247	Sequence 35	from Patent
gb_pat:AX217532	17	16.00	101.77	2.6e+03	17	AX217532	Sequence 2973	from patent	gb_pat:AR029157	19	16.00	101.03	2.9e+03	19	AR029157	Sequence 33	from pat
gb_pat:AX217533	17	16.00	101.77	2.6e+03	17	AX217533	Sequence 2974	from patent	gb_pat:AR036541	19	16.00	101.03	2.9e+03	19	AR036541	Sequence 33	from pat
gb_pat:AX218159	17	16.00	101.77	2.6e+03	17	AX218159	Sequence 2975	from patent	gb_pat:AR096074	19	16.00	101.03	2.9e+03	19	AR096074	Sequence 33	from pat
gb_pat:AX218200	17	16.00	101.77	2.6e+03	17	AX218200	Sequence 3601	from patent	gb_pat:AR102546	19	16.00	101.03	2.9e+03	19	AR102546	Sequence 35	from pat
gb_pat:AX218290	17	16.00	101.77	2.6e+03	17	AX218290	Sequence 3642	from patent	gb_pat:AR111930	19	16.00	101.03	2.9e+03	19	AR111930	Sequence 4	from pat
gb_pat:AX227570	17	16.00	101.77	2.6e+03	17	AX227570	Sequence 9432	from patent	gb_pat:AR124827	19	16.00	101.03	2.9e+03	19	AR124827	Sequence 4	from pat
gb_pat:AX263363	17	16.00	101.77	2.6e+03	17	AX263363	Sequence 755	from patent	gb_pat:AR135275	19	16.00	101.03	2.9e+03	19	AR135275	Sequence 4	from pat
gb_pat:AX263365	17	16.00	101.77	2.6e+03	17	AX263365	Sequence 755	from patent	gb_pat:AR141345	19	16.00	101.03	2.9e+03	19	AR141345	Sequence 12	from pat
gb_pat:AX263596	17	16.00	101.77	2.6e+03	17	AX263596	Sequence 756	from patent	gb_pat:AX131071	19	16.00	101.03	2.9e+03	19	AX131071	Sequence 2289	from P
gb_pat:AX263597	17	16.00	101.77	2.6e+03	17	AX263597	Sequence 987	from patent	gb_pat:AX131072	19	16.00	101.03	2.9e+03	19	AX131072	Sequence 2290	from P
gb_pat:AX264244	17	16.00	101.77	2.6e+03	17	AX264244	Sequence 988	from patent	gb_pat:AX131073	19	16.00	101.03	2.9e+03	19	AX131073	Sequence 2291	from P
gb_pat:AX264245	17	16.00	101.77	2.6e+03	17	AX264245	Sequence 1635	from patent	gb_pat:AX131074	19	16.00	101.03	2.9e+03	19	AX131074	Sequence 2292	from P
gb_pat:BD007682	17	16.00	101.77	2.6e+03	17	BD007682	Sequence 1636	from patent	gb_pat:AX131075	19	16.00	101.03	2.9e+03	19	AX131075	Sequence 2293	from P
gb_pat:BD007683	17	16.00	101.77	2.6e+03	17	BD007683	Sequence 1636	from patent	gb_pat:AX131076	19	16.00	101.03	2.9e+03	19	AX131076	Sequence 2294	from P
gb_pat:BD007684	17	16.00	101.77	2.6e+03	17	BD007684	Sequence 1636	from patent	gb_pat:AX131077	19	16.00	101.03	2.9e+03	19	AX131077	Sequence 2295	from P
gb_pat:E08103	17	16.00	101.77	2.6e+03	17	E08103	Sequence 597	from patent	gb_pat:AX131078	19	16.00	101.03	2.9e+03	19	AX131078	Sequence 2296	from P
gb_pat:E34258	17	16.00	101.77	2.6e+03	17	E34258	Synthetic oligonucleotide		gb_pat:AX131079	19	16.00	101.03	2.9e+03	19	AX131079	Sequence 3191	from P
gb_pat:E34259	17	16.00	101.77	2.6e+03	17	E34259	Pollinosis-associated ger		gb_pat:AX131074	19	16.00	101.03	2.9e+03	19	AX131074	Sequence 3192	from P
gb_pat:E34260	17	16.00	101.77	2.6e+03	17	E34260	Pollinosis-associated ger		gb_pat:AX131075	19	16.00	101.03	2.9e+03	19	AX131075	Sequence 3193	from P
gb_pat:E59657	17	16.00	101.77	2.6e+03	17	E59657	Pollinosis-associated ger		gb_pat:AX131076	19	16.00	101.03	2.9e+03	19	AX131076	Sequence 3194	from P
gb_pat:137584	17	16.00	101.77	2.6e+03	17	137584	Method for preparing nucl		gb_pat:AX131077	19	16.00	101.03	2.9e+03	19	AX131077	Sequence 3195	from P
gb_pat:153049	17	16.00	101.77	2.6e+03	17	153049	Sequence 597	from patent	gb_pat:AX131078	19	16.00	101.03	2.9e+03	19	AX131078	Sequence 3196	from P
gb_pat:153325	17	16.00	101.77	2.6e+03	17	153325	Sequence 790	from patent	gb_pat:AX131079	19	16.00	101.03	2.9e+03	19	AX131079	Sequence 3197	from P
gb_pat:153327	17	16.00	101.77	2.6e+03	17	153327	Sequence 1066	from patent	gb_pat:AX132830	19	16.00	101.03	2.9e+03	19	AX132830	Sequence 4047	from P
gb_pat:153329	17	16.00	101.77	2.6e+03	17	153329	Sequence 1068	from patent	gb_pat:AX132831	19	16.00	101.03	2.9e+03	19	AX132831	Sequence 4048	from P
gb_pat:153331	17	16.00	101.77	2.6e+03	17	153331	Sequence 1070	from patent	gb_pat:AX132832	19	16.00	101.03	2.9e+03	19	AX132832	Sequence 4050	from P
gb_pat:153333	17	16.00	101.77	2.6e+03	17	153333	Sequence 1072	from patent	gb_pat:AX132833	19	16.00	101.03	2.9e+03	19	AX132833	Sequence 4051	from P
gb_pat:153335	17	16.00	101.77	2.6e+03	17	153335	Sequence 1074	from patent	gb_pat:AX132834	19	16.00	101.03	2.9e+03	19	AX132834	Sequence 4052	from P
gb_pat:153337	17	16.00	101.77	2.6e+03	17	153337	Sequence 1076	from patent	gb_pat:AX132835	19	16.00	101.03	2				



gb_pat:AX352893	19	1	AX352893	Sequence 99 from Patent	16.00	101.03	2.9e+03	gb_pat:AX136904	20	1	AX136904	Sequence 6 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352895	19	1	AX352895	Sequence 101 from Patent	16.00	101.03	2.9e+03	gb_pat:AX148814	20	1	AX148814	Sequence 16 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352902	19	1	AX352902	Sequence 108 from Patent	16.00	101.03	2.9e+03	gb_pat:AX148971	20	1	AX148971	Sequence 173 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352904	19	1	AX352904	Sequence 110 from Patent	16.00	101.03	2.9e+03	gb_pat:AX167857	20	1	AX167857	Sequence 41 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352911	19	1	AX352911	Sequence 117 from Patent	16.00	101.03	2.9e+03	gb_pat:AX167928	20	1	AX167928	Sequence 112 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352913	19	1	AX352913	Sequence 119 from Patent	16.00	101.03	2.9e+03	gb_pat:AX167929	20	1	AX167929	Sequence 113 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352922	19	1	AX352922	Sequence 128 from Patent	16.00	101.03	2.9e+03	gb_pat:AX167930	20	1	AX167930	Sequence 114 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352923	19	1	AX352923	Sequence 129 from Patent	16.00	101.03	2.9e+03	gb_pat:AX250615	20	1	AX250615	Sequence 11 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352925	19	1	AX352925	Sequence 131 from Patent	16.00	101.03	2.9e+03	gb_pat:AX294371	20	1	AX294371	Sequence 6133 from P	16.00	100.68	3.0e+03
gb_pat:AX352927	19	1	AX352927	Sequence 133 from Patent	16.00	101.03	2.9e+03	gb_pat:AX353136	20	1	AX353136	Sequence 184 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352933	19	1	AX352933	Sequence 139 from Patent	16.00	101.03	2.9e+03	gb_pat:AX355404	20	1	AX355404	Sequence 432 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352937	19	1	AX352937	Sequence 143 from Patent	16.00	101.03	2.9e+03	gb_pat:AX355709	20	1	AX355709	Sequence 737 from Pat	16.00	100.68	3.0e+03
gb_pat:AX352943	19	1	AX352943	Sequence 149 from Patent	16.00	101.03	2.9e+03	gb_pat:BD001838	20	1	BD001838	Method for identifi	16.00	100.68	3.0e+03
gb_pat:AX353097	19	1	AX353097	Sequence 303 from Patent	16.00	101.03	2.9e+03	gb_pat:BD008240	20	1	BD008240	Gene coding for prot	16.00	100.68	3.0e+03
gb_pat:BD008242	19	1	BD008242	Gene coding for protein	16.00	101.03	2.9e+03	gb_pat:E05831	20	1	E05831	DNA sequence for synth	16.00	100.68	3.0e+03
gb_pat:E29790	19	1	E29790	Method for discriminating	16.00	101.03	2.9e+03	gb_pat:E06910	20	1	E06910	Synthetic DNA for hydr	16.00	100.68	3.0e+03
gb_pat:I34186	19	1	I34186	Sequence 2 from patent US	16.00	101.03	2.9e+03	gb_pat:E29051	20	1	E29051	Mutant S182(Ps-1) gene	16.00	100.68	3.0e+03
gb_pat:AI01788	20	1	AI01788	Oligonucleotide 1 from pa	16.00	100.68	3.0e+03	gb_pat:E29055	20	1	E29055	Mutant S182(Ps-1) gene	16.00	100.68	3.0e+03
gb_pat:AI01129	20	1	AI01129	Sequence 1 from Patent WC	16.00	100.68	3.0e+03	gb_pat:E29057	20	1	E29057	Mutant S182(Ps-1) gene	16.00	100.68	3.0e+03
gb_pat:AI01887	20	1	AI01887	Sequence 5 from Patent WC	16.00	100.68	3.0e+03	gb_pat:E29065	20	1	E29065	Mutant S182(Ps-1) gene	16.00	100.68	3.0e+03
gb_pat:AR000176	20	1	AR000176	Sequence 7 from Patent	16.00	100.68	3.0e+03	gb_pat:E33607	20	1	E33607	Novel prokaryotic poly	16.00	100.68	3.0e+03
gb_pat:AR016001	20	1	AR016001	Sequence 21 from Patent	16.00	100.68	3.0e+03	gb_pat:E36060	20	1	E36060	Higher-order structure	16.00	100.68	3.0e+03
gb_pat:AR026504	20	1	AR026504	Sequence 11 from Patent	16.00	100.68	3.0e+03	gb_pat:E36060	20	1	E36060	Higher-order structure	16.00	100.68	3.0e+03
gb_pat:AR0242393	20	1	AR0242393	Sequence 13 from Patent	16.00	100.68	3.0e+03	gb_pat:E36664	20	1	E36664	DNA and plasmid contai	16.00	100.68	3.0e+03
gb_pat:AR050226	20	1	AR050226	Sequence 32 from Patent	16.00	100.68	3.0e+03	gb_pat:E51768	20	1	E51768	Promoter of melon ethy	16.00	100.68	3.0e+03
gb_pat:AR062036	20	1	AR062036	Sequence 115 from Patent	16.00	100.68	3.0e+03	gb_pat:E58612	20	1	E58612	Thermotolerant esteras	16.00	100.68	3.0e+03
gb_pat:AR082042	20	1	AR082042	Sequence 21 from Patent	16.00	100.68	3.0e+03	gb_pat:E59328	20	1	E59328	Method for purifying o	16.00	100.68	3.0e+03
gb_pat:AR082552	20	1	AR082552	Sequence 2 from Patent	16.00	100.68	3.0e+03	gb_pat:I16848	20	1	I16848	Sequence 13 from Patent	16.00	100.68	3.0e+03
gb_pat:AR086161	20	1	AR086161	Sequence 51 from Patent	16.00	100.68	3.0e+03	gb_pat:I23825	20	1	I23825	Sequence 4 from Patent	16.00	100.68	3.0e+03
gb_pat:AR087161	20	1	AR087161	Sequence 31 from Patent	16.00	100.68	3.0e+03	gb_pat:I25376	20	1	I25376	Sequence 13 from Patent	16.00	100.68	3.0e+03
gb_pat:AR087161	20	1	AR087161	Sequence 31 from Patent	16.00	100.68	3.0e+03	gb_pat:I36544	20	1	I36544	Sequence 13 from Patent	16.00	100.68	3.0e+03
gb_pat:AR089398	20	1	AR089398	Sequence 157 from Patent	16.00	100.68	3.0e+03	gb_pat:I49616	20	1	I49616	Sequence 31 from Patent	16.00	100.68	3.0e+03
gb_pat:AR089580	20	1	AR089580	Sequence 4 from Patent	16.00	100.68	3.0e+03	gb_pat:I49616	20	1	I49616	Sequence 31 from Patent	16.00	100.68	3.0e+03
gb_pat:AR092979	20	1	AR092979	Sequence 74 from Patent	16.00	100.68	3.0e+03	gb_sy:AB069479	20	1	AB069479	Synthetic construct	16.00	100.68	3.0e+03
gb_pat:AR093063	20	1	AR093063	Sequence 158 from Patent	16.00	100.68	3.0e+03	em_pat:E09534	20	1	E09534	Synthetic DNA. 9/2000	16.00	100.68	3.0e+03
gb_pat:AR093598	20	1	AR093598	Sequence 157 from Patent	16.00	100.68	3.0e+03	em_pat:E09785	20	1	E09785	DNA fragment for ligat	16.00	100.68	3.0e+03
gb_pat:AR093676	20	1	AR093676	Sequence 51 from Patent	16.00	100.68	3.0e+03	gb_om:DOGP2101	20	1	DOGP2101	Clone: CXX.21 pr	16.00	100.35	3.2e+03
gb_pat:AR096477	20	1	AR096477	Sequence 6 from Patent	16.00	100.68	3.0e+03	gb_om:DOGP41401	20	1	DOGP41401	Clone: CXX.414 p	16.00	100.35	3.2e+03
gb_pat:AR098294	20	1	AR098294	Sequence 4 from Patent	16.00	100.68	3.0e+03	gb_om:DOGP448902	20	1	DOGP448902	Clone: CXX.489 p	16.00	100.35	3.2e+03
gb_pat:AR100615	20	1	AR100615	Sequence 11 from Patent	16.00	100.68	3.0e+03	gb_pat:A23687	20	1	A23687	L. monocytogenes HlyA	16.00	100.35	3.2e+03
gb_pat:AR101034	20	1	AR101034	Sequence 4 from Patent	16.00	100.68	3.0e+03	gb_pat:AX5600	20	1	AX5600	Sequence 5 from Patent	16.00	100.35	3.2e+03
gb_pat:AR101040	20	1	AR101040	Sequence 10 from Patent	16.00	100.68	3.0e+03	gb_pat:AX5600	20	1	AX5600	Sequence 5 from Patent	16.00	100.35	3.2e+03
gb_pat:AR101049	20	1	AR101049	Sequence 19 from Patent	16.00	100.68	3.0e+03	gb_pat:AX83596	20	1	AX83596	Sequence 25 from Patent	16.00	100.35	3.2e+03
gb_pat:AR105641	20	1	AR105641	Sequence 3 from Patent	16.00	100.68	3.0e+03	gb_pat:AX91562	20	1	AX91562	Sequence 89 from Patent	16.00	100.35	3.2e+03
gb_pat:AR118958	20	1	AR118958	Sequence 84 from Patent	16.00	100.68	3.0e+03	gb_pat:AR000177	20	1	AR000177	Sequence 8 from Patent	16.00	100.35	3.2e+03
gb_pat:AR118959	20	1	AR118959	Sequence 85 from Patent	16.00	100.68	3.0e+03	gb_pat:AR016251	20	1	AR016251	Sequence 5 from Patent	16.00	100.35	3.2e+03
gb_pat:AR121076	20	1	AR121076	Sequence 97 from Patent	16.00	100.68	3.0e+03	gb_pat:AR060940	20	1	AR060940	Sequence 288 from Pat	16.00	100.35	3.2e+03
gb_pat:AR121615	20	1	AR121615	Sequence 12 from Patent	16.00	100.68	3.0e+03	gb_pat:AR050526	20	1	AR050526	Sequence 3 from Patent	16.00	100.35	3.2e+03
gb_pat:AR122236	20	1	AR122236	Sequence 82 from Patent	16.00	100.68	3.0e+03	gb_pat:AR103576	20	1	AR103576	Sequence 100 from Pat	16.00	100.35	3.2e+03
gb_pat:AR122473	20	1	AR122473	Sequence 27 from Patent	16.00	100.68	3.0e+03	gb_pat:AR103576	20	1	AR103576	Sequence 23 from Pat	16.00	100.35	3.2e+03
gb_pat:AR131210	20	1	AR131210	Sequence 82 from Patent	16.00	100.68	3.0e+03	gb_pat:AR142678	20	1	AR142678	Sequence 8 from Patent	16.00	100.35	3.2e+03
gb_pat:AR142339	20	1	AR142339	Sequence 21 from Patent	16.00	100.68	3.0e+03	gb_pat:AR153849	20	1	AR153849	Sequence 2 from Patent	16.00	100.35	3.2e+03
gb_pat:AR142340	20	1	AR142340	Sequence 22 from Patent	16.00	100.68	3.0e+03	gb_pat:AX023419	20	1	AX023419	Sequence 34 from Pat	16.00	100.35	3.2e+03
gb_pat:AR142341	20	1	AR142341	Sequence 23 from Patent	16.00	100.68	3.0e+03	gb_pat:AX111720	20	1	AX111720	Sequence 17 from Pat	16.00	100.35	3.2e+03
gb_pat:AR142342	20	1	AR142342	Sequence 24 from Patent	16.00	100.68	3.0e+03	gb_pat:AX111723	20	1	AX111723	Sequence 20 from Pat	16.00	100.35	3.2e+03
gb_pat:AR143169	20	1	AR143169	Sequence 62 from Patent	16.00	100.68	3.0e+03	gb_pat:AX111910	20	1	AX111910	Sequence 18 from Pat	16.00	100.35	3.2e+03
gb_pat:AR144624	20	1	AR144624	Sequence 169 from Patent	16.00	100.68	3.0e+03	gb_pat:AX119911	20	1	AX119911	Sequence 19 from Pat	16.00	100.35	3.2e+03
gb_pat:AR149897	20	1	AR149897	Sequence 4 from Patent	16.00	100.68	3.0e+03	gb_pat:AX133303	20	1	AX133303	Sequence 4521 from P	16.00	100.35	3.2e+03
gb_pat:AR160701	20	1	AR160701	Sequence 88 from Patent	16.00	100.68	3.0e+03	gb_pat:AX154158	20	1	AX154158	Sequence 286 from Pat	16.00	100.35	3.2e+03
gb_pat:AR160702	20	1	AR160702	Sequence 89 from Patent	16.00	100.68	3.0e+03	gb_pat:AX288083	20	1	AX288083	Sequence 28 from Pat	16.00	100.35	3.2e+03
gb_pat:AR167073	20	1	AR167073	Sequence 90 from Patent	16.00	100.68	3.0e+03	gb_pat:E08958	20	1	E08958	Probe of Stb-gene. 9/1	16.00	100.35	3.2e+03
gb_pat:AR167074	20	1	AR167074	Sequence 91 from Patent	16.00	100.68	3.0e+03	gb_pat:E28097	20	1	E28097	Method for analyzing D	16.00	100.35	3.2e+03
gb_pat:AR167077	20	1	AR167077	Sequence 94 from Patent	16.00	100.68	3.0e+03	gb_pat:E29379	20	1	E29379	Oligonucleotide and re	16.00	100.35	3.2e+03
gb_pat:AR173877	20	1	AR173877	Sequence 75 from Patent	16.00	100.68	3.0e+03	gb_pat:I22181	20	1	I22181	Sequence 40 from Patent	16.00	100.35	3.2e+03
gb_pat:AR173878	20	1	AR173878	Sequence 76 from Patent	16.00	100.68	3.0e+03	gb_pat:I30723	20	1	I30723	Sequence 161 from Pat	16.00	100.35	3.2e+03
gb_pat:AX076462	20	1	AX076462	Sequence 7 from Patent	16.00	100.68	3.0e+03	gb_pat:I30769	20	1	I30769	Sequence 207 from Pat	16.00	100.35	3.2e+03
gb_pat:AX1042393	20	1	AX1042393	Sequence 431 from Patent	16.00	100.68	3.0e+03	gb_pat:I36166	20	1	I36166	Sequence 2 from Patent	16.00	100.35	3.2e+03
gb_pat:AX104365	20	1	AX104365	Sequence 557 from Patent	16.00	100.68	3.0e+03	gb_pat:I39871	20	1	I39871	Sequence 1 from Patent	16.00	100.35	3.2e+03
gb_pat:AX104369	20	1	AX104369	Sequence 561 from Patent	16.00	100.68	3.0e+03	gb_pat:I46182	20	1	I46182	Sequence 161 from Pat	16.00	100.35	3.2e+03
gb_pat:AX104899	20	1	AX104899	Sequence 1091 from Patent	16.00	100.68	3.0e+03	gb_pat:I46228	20	1	I46228	Sequence 207 from Pat	16.00	100.35	3.2e+03
gb_pat:AX109817	20	1	AX109817	Sequence 550 from Patent	16.00	100.68	3.0e+03	gb_pat:I56671	20	1	I56671	Sequence 2 from Patent	16.00	100.35	3.2e+03
gb_pat:AX109867	20	1	AX109867	Sequence 600 from											



gb_in:DME421765	-	16.00	100.04	3.3e+03	22	AJ421765 Drosophila melanogaster	gb_pat:I04013	16.00	99.74	3.4e+03	23	I04013 Sequence 9 from Patent
gb_in:DME421766	+	16.00	100.04	3.3e+03	22	AJ421766 Drosophila melanogaster	gb_pat:I26414	16.00	99.74	3.4e+03	23	I26414 Sequence 106 from patent
gb_pat:A42091	+	16.00	100.04	3.3e+03	22	A42091 Sequence 3 from Patent WO	gb_pat:I30437	16.00	99.74	3.4e+03	23	I30437 Sequence 24 from patent
gb_pat:A51385	+	16.00	100.04	3.3e+03	22	A51385 Sequence 1 from Patent WO	gb_pat:I32906	16.00	99.74	3.4e+03	23	I32906 Sequence 15 from patent
gb_pat:A51391	+	16.00	100.04	3.3e+03	22	A51391 Sequence 7 from Patent WO	gb_pat:I35681	16.00	99.74	3.4e+03	23	I35681 Sequence 24 from patent
gb_pat:A85577	-	16.00	100.04	3.3e+03	22	A85577 Sequence 6 from Patent WO	gb_pat:I79498	16.00	99.74	3.4e+03	23	I79498 Sequence 5 from patent
gb_pat:AR072065	+	16.00	100.04	3.3e+03	22	AR072065 Sequence 1 from patent	gb_pat:I79499	16.00	99.74	3.4e+03	23	I79499 Sequence 6 from patent
gb_pat:AR092325	+	16.00	100.04	3.3e+03	22	AR092325 Sequence 11 from patent	gb_pat:I96223	16.00	99.74	3.4e+03	23	I96223 Sequence 5 from patent
gb_pat:AR092326	+	16.00	100.04	3.3e+03	22	AR092326 Sequence 12 from patent	gb_pat:I96223	16.00	99.74	3.4e+03	23	I96223 Sequence 5 from patent
gb_pat:AR094982	+	16.00	100.04	3.3e+03	22	AR094982 Sequence 20 from patent	gb_pr:S61964	16.00	99.74	3.4e+03	23	S61964 (beta C)-beta-globin (
gb_pat:AR164318	+	16.00	100.04	3.3e+03	22	AR164318 Sequence 1 from patent	gb_pr:S61969	16.00	99.74	3.4e+03	23	S61969 (A3 beta C)-beta-globi
gb_pat:AR164319	+	16.00	100.04	3.3e+03	22	AR164319 Sequence 2 from patent	gb_sts:HUMSTS21LF	16.00	99.74	3.4e+03	23	HUMSTS21LF Human SPS primer ppw51
gb_pat:AR176636	+	16.00	100.04	3.3e+03	22	AR176636 Sequence 3 from patent	gb_sts:HUMSTS21LF	16.00	99.74	3.4e+03	23	HUMSTS21LF Human SPS primer ppw51
gb_pat:AX105723	+	16.00	100.04	3.3e+03	22	AX105723 Sequence 26 from patent	gb_pat:A20374	16.00	99.46	3.5e+03	24	A20374 Oligonucleotide. 10/19
gb_pat:AX115334	+	16.00	100.04	3.3e+03	22	AX115334 Sequence 26 from Patent	gb_pat:A73476	16.00	99.46	3.5e+03	24	A73476 Synthetic pcAS7 poliov
gb_pat:AX116999	+	16.00	100.04	3.3e+03	22	AX116999 Sequence 457 from Patent	gb_pat:A74269	16.00	99.46	3.5e+03	24	A74269 Synthetic 25 from Patent
gb_pat:AX1196214	+	16.00	100.04	3.3e+03	22	AX1196214 Sequence 2122 from Patent	gb_pat:AR010030	16.00	99.46	3.5e+03	24	AR010030 Sequence 43 from pat
gb_pat:AX1196216	+	16.00	100.04	3.3e+03	22	AX1196216 Sequence 45 from Patent	gb_pat:AR010031	16.00	99.46	3.5e+03	24	AR010031 Sequence 44 from pat
gb_pat:AX351956	+	16.00	100.04	3.3e+03	22	AX351956 Sequence 252 from Patent	gb_pat:AR010032	16.00	99.46	3.5e+03	24	AR010032 Sequence 45 from pat
gb_pat:AX352052	-	16.00	100.04	3.3e+03	22	AX352052 Sequence 348 from Patent	gb_pat:AR010033	16.00	99.46	3.5e+03	24	AR010033 Sequence 46 from pat
gb_pat:AX352315	-	16.00	100.04	3.3e+03	22	AX352315 Sequence 611 from Patent	gb_pat:AR010034	16.00	99.46	3.5e+03	24	AR010034 Sequence 47 from pat
gb_pat:EX26475	+	16.00	100.04	3.3e+03	22	EX26475 Erysielothrix rhusiopath	gb_pat:AR010035	16.00	99.46	3.5e+03	24	AR010035 Sequence 48 from pat
gb_pat:E36534	+	16.00	100.04	3.3e+03	22	E36534 Method of gene diagnosis	gb_pat:AR019570	16.00	99.46	3.5e+03	24	AR019570 Sequence 55 from pat
gb_pat:E40156	+	16.00	100.04	3.3e+03	22	E40156 Genetic diagnosis method	gb_pat:AR034765	16.00	99.46	3.5e+03	24	AR034765 Sequence 43 from pat
gb_pat:I17096	+	16.00	100.04	3.3e+03	22	I17096 Sequence 11 from patent	gb_pat:AR034766	16.00	99.46	3.5e+03	24	AR034766 Sequence 44 from pat
gb_pat:I22660	+	16.00	100.04	3.3e+03	22	I22660 Sequence 148 from patent	gb_pat:AR034767	16.00	99.46	3.5e+03	24	AR034767 Sequence 45 from pat
gb_pat:I22661	+	16.00	100.04	3.3e+03	22	I22661 Sequence 149 from patent	gb_pat:AR034768	16.00	99.46	3.5e+03	24	AR034768 Sequence 46 from pat
gb_pat:I31810	+	16.00	100.04	3.3e+03	22	I31810 Sequence 1 from patent	gb_pat:AR034770	16.00	99.46	3.5e+03	24	AR034770 Sequence 48 from pat
gb_pat:I31811	+	16.00	100.04	3.3e+03	22	I31811 Sequence 2 from patent	gb_pat:AR058189	16.00	99.46	3.5e+03	24	AR058189 Sequence 5 from pat
gb_pat:I47485	+	16.00	100.04	3.3e+03	22	I47485 Sequence 148 from patent	gb_pat:AR058881	16.00	99.46	3.5e+03	24	AR058881 Sequence 13 from pat
gb_pat:I47486	+	16.00	100.04	3.3e+03	22	I47486 Sequence 149 from patent	gb_pat:AR079586	16.00	99.46	3.5e+03	24	AR079586 Sequence 10 from pat
gb_pat:I63491	+	16.00	100.04	3.3e+03	22	I63491 Sequence 11 from patent	gb_pat:AR081684	16.00	99.46	3.5e+03	24	AR081684 Sequence 11 from pat
gb_pat:I65356	+	16.00	100.04	3.3e+03	22	I65356 Sequence 5 from patent	gb_pat:AR084224	16.00	99.46	3.5e+03	24	AR084224 Sequence 11 from pat
gb_pat:I69407	+	16.00	100.04	3.3e+03	22	I69407 Sequence 1 from patent	gb_pat:AR109531	16.00	99.46	3.5e+03	24	AR109531 Sequence 55 from pat
gb_pat:I69408	+	16.00	100.04	3.3e+03	22	I69408 Sequence 2 from patent	gb_pat:AR123295	16.00	99.46	3.5e+03	24	AR123295 Sequence 13 from pat
gb_sy:AR069603	-	16.00	100.04	3.3e+03	22	AR069603 Synthetic construct DNA	gb_pat:AR141034	16.00	99.46	3.5e+03	24	AR141034 Sequence 10 from pat
gb_pat:AR491179	-	16.00	99.74	3.4e+03	23	A491179 Sequence 13 from Patent W	gb_pat:AR162082	16.00	99.46	3.5e+03	24	AR162082 Sequence 8 from pat
gb_pat:AR014285	+	16.00	99.74	3.4e+03	23	AR014285 Sequence 17 from patent	gb_pat:AR166607	16.00	99.46	3.5e+03	24	AR166607 Sequence 10 from pat
gb_pat:AR016547	+	16.00	99.74	3.4e+03	23	AR016547 Sequence 24 from patent	gb_pat:AR168453	16.00	99.46	3.5e+03	24	AR168453 Sequence 82 from pat
gb_pat:AR021267	+	16.00	99.74	3.4e+03	23	AR021267 Sequence 24 from patent	gb_pat:AR093355	16.00	99.46	3.5e+03	24	AR093355 Sequence 85 from Pat
gb_pat:AR036821	+	16.00	99.74	3.4e+03	23	AR036821 Sequence 24 from patent	gb_pat:AX104753	16.00	99.46	3.5e+03	24	AX104753 Sequence 945 from Pa
gb_pat:AR038118	+	16.00	99.74	3.4e+03	23	AR038118 Sequence 23 from patent	gb_pat:AX117090	16.00	99.46	3.5e+03	24	AX117090 Sequence 2213 from P
gb_pat:AR038190	+	16.00	99.74	3.4e+03	23	AR038190 Sequence 6 from patent	gb_pat:AX184168	16.00	99.46	3.5e+03	24	AX184168 Sequence 1921 from P
gb_pat:AR061895	+	16.00	99.74	3.4e+03	23	AR061895 Sequence 26 from patent	gb_pat:AX2899738	16.00	99.46	3.5e+03	24	AX2899738 Sequence 1500 from P
gb_pat:AR066332	+	16.00	99.74	3.4e+03	23	AR066332 Sequence 6 from patent	gb_pat:E32335	16.00	99.46	3.5e+03	24	E32335 Species-specific detec
gb_pat:AR072303	+	16.00	99.74	3.4e+03	23	AR072303 Sequence 106 from patent	gb_pat:E36596	16.00	99.46	3.5e+03	24	E36596 Antisense nucleic acid
gb_pat:AR082880	+	16.00	99.74	3.4e+03	23	AR082880 Sequence 13 from patent	gb_pat:I06840	16.00	99.46	3.5e+03	24	I06840 Sequence 3 from Patent
gb_pat:AR084141	+	16.00	99.74	3.4e+03	23	AR084141 Sequence 44 from patent	gb_pat:I06842	16.00	99.46	3.5e+03	24	I06842 Sequence 5 from Patent
gb_pat:AR084981	+	16.00	99.74	3.4e+03	23	AR084981 Sequence 15 from patent	gb_pat:I24741	16.00	99.46	3.5e+03	24	I24741 Sequence 4 from patent
gb_pat:AR091436	+	16.00	99.74	3.4e+03	23	AR091436 Sequence 5 from patent	gb_pat:I24742	16.00	99.46	3.5e+03	24	I24742 Sequence 5 from patent
gb_pat:AR091439	+	16.00	99.74	3.4e+03	23	AR091439 Sequence 14 from patent	gb_pat:I24743	16.00	99.46	3.5e+03	24	I24743 Sequence 6 from patent
gb_pat:AR095431	+	16.00	99.74	3.4e+03	23	AR095431 Sequence 9 from patent	gb_pat:I24746	16.00	99.46	3.5e+03	24	I24746 Sequence 9 from patent
gb_pat:AR096943	+	16.00	99.74	3.4e+03	23	AR096943 Sequence 78 from patent	gb_pat:I24747	16.00	99.46	3.5e+03	24	I24747 Sequence 10 from patent
gb_pat:AR097321	+	16.00	99.74	3.4e+03	23	AR097321 Sequence 8 from patent	gb_pat:I24748	16.00	99.46	3.5e+03	24	I24748 Sequence 11 from patent
gb_pat:AR103040	+	16.00	99.74	3.4e+03	23	AR103040 Sequence 17 from patent	gb_pat:I32528	16.00	99.46	3.5e+03	24	I32528 Sequence 13 from patent
gb_pat:AR123791	+	16.00	99.74	3.4e+03	23	AR123791 Sequence 7 from patent	gb_pat:I35523	16.00	99.46	3.5e+03	24	I35523 Sequence 13 from patent
gb_pat:AR138953	+	16.00	99.74	3.4e+03	23	AR138953 Sequence 5 from patent	gb_pat:I43133	16.00	99.46	3.5e+03	24	I43133 Sequence 13 from patent
gb_pat:AR138956	+	16.00	99.74	3.4e+03	23	AR138956 Sequence 14 from patent	gb_pat:I451702	16.00	99.46	3.5e+03	24	I451702 Sequence 55 from patent
gb_pat:AR145827	+	16.00	99.74	3.4e+03	23	AR145827 Sequence 6 from patent	gb_pat:I65357	16.00	99.46	3.5e+03	24	I65357 Sequence 6 from patent
gb_pat:AR164605	+	16.00	99.74	3.4e+03	23	AR164605 Sequence 5 from patent	gb_pat:I76479	16.00	99.46	3.5e+03	24	I76479 Sequence 55 from patent
gb_pat:AR164655	+	16.00	99.74	3.4e+03	23	AR164655 Sequence 10 from patent	gb_pat:I90367	16.00	99.46	3.5e+03	24	I90367 Sequence 11 from patent
gb_pat:AX116398	+	16.00	99.74	3.4e+03	23	AX116398 Sequence 1521 from Patent	gb_pat:I92011	16.00	99.46	3.5e+03	24	I92011 Sequence 13 from patent
gb_pat:AX118499	+	16.00	99.74	3.4e+03	23	AX118499 Sequence 3622 from Patent	gb_ro:MUSBMP24B	16.00	99.46	3.5e+03	24	BMP-2/4- Mouse gene for BMP-2/4
gb_pat:AX183613	+	16.00	99.74	3.4e+03	23	AX183613 Sequence 1366 from Patent	gb_sy:DOGAPOBA	16.00	99.46	3.5e+03	24	DOGAPOBA Canis familiaris Apol1
gb_pat:AX250620	+	16.00	99.74	3.4e+03	23	AX250620 Sequence 16 from Patent	gb_sy:AB068635	16.00	99.46	3.5e+03	24	AB068635 Synthetic construct
gb_pat:AX269348	+	16.00	99.74	3.4e+03	23	AX269348 Sequence 6 from Patent	gb_pat:AX12768	16.00	99.19	3.7e+03	25	AX12768 Synthetic promoter and
gb_pat:E12392	+	16.00	99.74	3.4e+03	23	E12392 Oligonucleotide primer. 6	gb_pat:AX18807	16.00	99.19	3.7e+03	25	AX18807 Oligonucleotide primer
gb_pat:E12393	+	16.00	99.74	3.4e+03	23	E12393 Oligonucleotide primer. 6	gb_pat:AX62521	16.00	99.19	3.7e+03	25	AX62521 Sequence 26 from Patent
gb_pat:E12772	+	16.00	99.74	3.4e+03	23	E12772 PCR primer for gaining di	gb_pat:AX66944	16.00	99.19	3.7e+03	25	AX66944 Sequence 111 from Paten
gb_pat:E14519	+	16.00	99.74	3.4e+03	23	E14519 PCR primer for gaining Br	gb_pat:AX72394	16.00	99.19	3.7e+03	25	AX72394 Sequence 26 from Paten
gb_pat:E16748	+	16.00	99.74	3.4e+03	23	E16748 Primer. 7/1999	gb_pat:AR014025	16.00	99.19	3.7e+03	25	AR014025 Sequence 61 from pat
gb_pat:I00179	+	16.00	99.74	3.4e+03	23	I00179 Sequence 3 from Patent US	gb_pat:AR048532	16.00	99.19	3.7e+03	25	AR048532 Sequence 18 from pat
gb_pat:I00269	+	16.00	99.74	3.4e+03	23	I00269 Sequence 2 from Patent US	gb_pat:AR095203	16.00	99.19	3.7e+03	25	AR095203 Sequence 1 from pat



gb_pat:AR125170	-	16.00	99.19	3.7e+03	25	! AR125170 Sequence 1 from Patent	gb_pat:IL17013	+	16.00	98.92	3.8e+03	26	! IL17013 Sequence 24 from Patent
gb_pat:AR137554	+	16.00	99.19	3.7e+03	25	! AR137554 Sequence 39 from Patent	gb_pat:IL24758	-	16.00	98.92	3.8e+03	26	! IL24758 Sequence 21 from Patent
gb_pat:AR139060	+	16.00	99.19	3.7e+03	25	! AR139060 Sequence 39 from Patent	gb_pat:IL34036	-	16.00	98.92	3.8e+03	26	! IL34036 Sequence 11 from Patent
gb_pat:ARX003409	+	16.00	99.19	3.7e+03	25	! ARX003409 Sequence 26 from Patent	gb_pat:IL34052	-	16.00	98.92	3.8e+03	26	! IL34052 Sequence 29 from Patent
gb_pat:ARX007110	+	16.00	99.19	3.7e+03	25	! ARX007110 Sequence 10 from Patent	gb_pat:IL46641	+	16.00	98.92	3.8e+03	26	! IL46641 Sequence 620 from Patent
gb_pat:ARX03469	-	16.00	99.19	3.7e+03	25	! ARX03469 Sequence 34 from Patent	gb_pat:IL60424	-	16.00	98.92	3.8e+03	26	! IL60424 Sequence 11 from Patent
gb_pat:ARX104751	-	16.00	99.19	3.7e+03	25	! ARX104751 Sequence 943 from Patent	gb_pat:IL81318	+	16.00	98.92	3.8e+03	26	! IL81318 Sequence 19 from Patent
gb_pat:ARX110892	-	16.00	99.19	3.7e+03	25	! ARX110892 Sequence 1625 from Patent	gb_pat:AR001598	+	16.00	98.67	3.9e+03	27	! AR001598 Sequence 17 from Patent
gb_pat:ARX115660	-	16.00	99.19	3.7e+03	25	! ARX115660 Sequence 783 from Patent	gb_pat:AR039572	+	16.00	98.67	3.9e+03	27	! AR039572 Sequence 420 from Patent
gb_pat:ARX115988	-	16.00	99.19	3.7e+03	25	! ARX115988 Sequence 1111 from Patent	gb_pat:AR066366	+	16.00	98.67	3.9e+03	27	! AR066366 Sequence 5 from Patent
gb_pat:ARX116004	-	16.00	99.19	3.7e+03	25	! ARX116004 Sequence 1127 from Patent	gb_pat:AR094979	+	16.00	98.67	3.9e+03	27	! AR094979 Sequence 17 from Patent
gb_pat:ARX116188	-	16.00	99.19	3.7e+03	25	! ARX116188 Sequence 1311 from Patent	gb_pat:AR094980	-	16.00	98.67	3.9e+03	27	! AR094980 Sequence 18 from Patent
gb_pat:ARX116396	+	16.00	99.19	3.7e+03	25	! ARX116396 Sequence 1519 from Patent	gb_pat:AR128034	+	16.00	98.67	3.9e+03	27	! AR128034 Sequence 184 from Patent
gb_pat:ARX116440	-	16.00	99.19	3.7e+03	25	! ARX116440 Sequence 1563 from Patent	gb_pat:AR176012	+	16.00	98.67	3.9e+03	27	! AR176012 Sequence 28 from Patent
gb_pat:ARX116671	+	16.00	99.19	3.7e+03	25	! ARX116671 Sequence 1794 from Patent	gb_pat:ARX035226	+	16.00	98.67	3.9e+03	27	! ARX035226 Sequence 2 from Patent
gb_pat:ARX116811	-	16.00	99.19	3.7e+03	25	! ARX116811 Sequence 1934 from Patent	gb_pat:ARX035528	+	16.00	98.67	3.9e+03	27	! ARX035528 Sequence 4 from Patent
gb_pat:ARX116840	-	16.00	99.19	3.7e+03	25	! ARX116840 Sequence 1963 from Patent	gb_pat:ARX039054	+	16.00	98.67	3.9e+03	27	! ARX039054 Sequence 7 from Patent
gb_pat:ARX117208	-	16.00	99.19	3.7e+03	25	! ARX117208 Sequence 2331 from Patent	gb_pat:ARX043901	+	16.00	98.67	3.9e+03	27	! ARX043901 Sequence 14 from Patent
gb_pat:ARX117788	-	16.00	99.19	3.7e+03	25	! ARX117788 Sequence 2911 from Patent	gb_pat:ARX055902	+	16.00	98.67	3.9e+03	27	! ARX055902 Sequence 38 from Patent
gb_pat:ARX118212	-	16.00	99.19	3.7e+03	25	! ARX118212 Sequence 3335 from Patent	gb_pat:ARX101032	+	16.00	98.67	3.9e+03	27	! ARX101032 Sequence 6 from Patent
gb_pat:ARX118336	-	16.00	99.19	3.7e+03	25	! ARX118336 Sequence 3459 from Patent	gb_pat:ARX115396	+	16.00	98.67	3.9e+03	27	! ARX115396 Sequence 519 from Patent
gb_pat:ARX151163	-	16.00	99.19	3.7e+03	25	! ARX151163 Sequence 52 from Patent	gb_pat:ARX115516	+	16.00	98.67	3.9e+03	27	! ARX115516 Sequence 639 from Patent
gb_pat:ARX155622	-	16.00	99.19	3.7e+03	25	! ARX155622 Sequence 34 from Patent	gb_pat:ARX118664	+	16.00	98.67	3.9e+03	27	! ARX118664 Sequence 1987 from Patent
gb_pat:ARX196867	-	16.00	99.19	3.7e+03	25	! ARX196867 Sequence 574 from Patent	gb_pat:ARX196210	+	16.00	98.67	3.9e+03	27	! ARX196210 Sequence 41 from Patent
gb_pat:ARX196988	-	16.00	99.19	3.7e+03	25	! ARX196988 Sequence 695 from Patent	gb_pat:ARX225178	+	16.00	98.67	3.9e+03	27	! ARX225178 Sequence 9 from Patent
gb_pat:ARX241161	+	16.00	99.19	3.7e+03	25	! ARX241161 Sequence 399 from Patent	gb_pat:ARX225180	+	16.00	98.67	3.9e+03	27	! ARX225180 Sequence 11 from Patent
gb_pat:IL13373	-	16.00	99.19	3.7e+03	25	! IL13373 Sequence 24 from Patent	gb_pat:ARX225182	+	16.00	98.67	3.9e+03	27	! ARX225182 Sequence 13 from Patent
gb_pat:IL21975	-	16.00	99.19	3.7e+03	25	! IL21975 Sequence 61 from Patent	gb_pat:ARX259894	+	16.00	98.67	3.9e+03	27	! ARX259894 Sequence 21 from Patent
gb_pat:IL45922	-	16.00	99.19	3.7e+03	25	! IL45922 Sequence 10 from Patent	gb_pat:ARX278516	+	16.00	98.67	3.9e+03	27	! ARX278516 Sequence 53 from Patent
gb_pr:HS231H41R	+	16.00	99.19	3.7e+03	25	! HS231H41R Sequence 1 of STS	gb_pat:ARX28769	+	16.00	98.67	3.9e+03	27	! ARX28769 Sequence 286 from Patent
gb_pr:HSMDAH1	+	16.00	99.19	3.7e+03	25	! HSMDAH1 Sequence 8 from Patent	gb_pat:IL15985	+	16.00	98.67	3.9e+03	27	! IL15985 Sequence 7 from Patent
gb_pat:AR24938	+	16.00	98.92	3.8e+03	26	! AR24938 Sequence 10 from Patent	gb_pat:IL58514	+	16.00	98.67	3.9e+03	27	! IL58514 Sequence 17 from Patent
gb_pat:AR9347	+	16.00	98.92	3.8e+03	26	! AR9347 Sequence 21 from Patent	gb_pat:IL63341	+	16.00	98.67	3.9e+03	27	! IL63341 Sequence 28 from Patent
gb_pat:AR010003	+	16.00	98.92	3.8e+03	26	! AR010003 Sequence 15 from Patent	gb_pr:S64862S3	+	16.00	98.67	3.9e+03	27	! S64864 alpha 1-theta 1 globin
gb_pat:AR017573	+	16.00	98.92	3.8e+03	26	! AR017573 Sequence 3 from Patent	gb_pat:AR0131	+	16.00	98.43	4.0e+03	28	! AR0131 Sequence 7 from Patent
gb_pat:AR017586	+	16.00	98.92	3.8e+03	26	! AR017586 Sequence 16 from Patent	gb_pat:AR014028	+	16.00	98.43	4.0e+03	28	! AR014028 Sequence 64 from Patent
gb_pat:AR020755	+	16.00	98.92	3.8e+03	26	! AR020755 Sequence 10 from Patent	gb_pat:AR014035	+	16.00	98.43	4.0e+03	28	! AR014035 Sequence 71 from Patent
gb_pat:AR020771	+	16.00	98.92	3.8e+03	26	! AR020771 Sequence 29 from Patent	gb_pat:AR022650	+	16.00	98.43	4.0e+03	28	! AR022650 Sequence 9 from Patent
gb_pat:AR034738	-	16.00	98.92	3.8e+03	26	! AR034738 Sequence 15 from Patent	gb_pat:AR055108	+	16.00	98.43	4.0e+03	28	! AR055108 Sequence 13 from Patent
gb_pat:AR038822	-	16.00	98.92	3.8e+03	26	! AR038822 Sequence 11 from Patent	gb_pat:AR055109	+	16.00	98.43	4.0e+03	28	! AR055109 Sequence 14 from Patent
gb_pat:AR098647	+	16.00	98.92	3.8e+03	26	! AR098647 Sequence 5 from Patent	gb_pat:AR055110	+	16.00	98.43	4.0e+03	28	! AR055110 Sequence 15 from Patent
gb_pat:AR136778	+	16.00	98.92	3.8e+03	26	! AR136778 Sequence 1 from Patent	gb_pat:AR055116	+	16.00	98.43	4.0e+03	28	! AR055116 Sequence 21 from Patent
gb_pat:AR144828	+	16.00	98.92	3.8e+03	26	! AR144828 Sequence 59 from Patent	gb_pat:AR055118	+	16.00	98.43	4.0e+03	28	! AR055118 Sequence 23 from Patent
gb_pat:ARX003453	+	16.00	98.92	3.8e+03	26	! ARX003453 Sequence 33 from Patent	gb_pat:AR068449	+	16.00	98.43	4.0e+03	28	! AR068449 Sequence 13 from Patent
gb_pat:ARX039875	+	16.00	98.92	3.8e+03	26	! ARX039875 Sequence 264 from Patent	gb_pat:AR068450	+	16.00	98.43	4.0e+03	28	! AR068450 Sequence 14 from Patent
gb_pat:ARX043897	+	16.00	98.92	3.8e+03	26	! ARX043897 Sequence 10 from Patent	gb_pat:AR068451	+	16.00	98.43	4.0e+03	28	! AR068451 Sequence 15 from Patent
gb_pat:ARX055900	+	16.00	98.92	3.8e+03	26	! ARX055900 Sequence 36 from Patent	gb_pat:AR068457	+	16.00	98.43	4.0e+03	28	! AR068457 Sequence 21 from Patent
gb_pat:ARX081473	-	16.00	98.92	3.8e+03	26	! ARX081473 Sequence 15 from Patent	gb_pat:AR068459	+	16.00	98.43	4.0e+03	28	! AR068459 Sequence 23 from Patent
gb_pat:ARX083734	-	16.00	98.92	3.8e+03	26	! ARX083734 Sequence 12 from Patent	gb_pat:AR068766	+	16.00	98.43	4.0e+03	28	! AR068766 Sequence 15 from Patent
gb_pat:ARX100410	-	16.00	98.92	3.8e+03	26	! ARX100410 Sequence 50 from Patent	gb_pat:AR085041	+	16.00	98.43	4.0e+03	28	! AR085041 Sequence 3 from Patent
gb_pat:ARX101041	+	16.00	98.92	3.8e+03	26	! ARX101041 Sequence 15 from Patent	gb_pat:AR085042	+	16.00	98.43	4.0e+03	28	! AR085042 Sequence 4 from Patent
gb_pat:ARX114990	+	16.00	98.92	3.8e+03	26	! ARX114990 Sequence 113 from Patent	gb_pat:AR085043	+	16.00	98.43	4.0e+03	28	! AR085043 Sequence 5 from Patent
gb_pat:ARX116767	+	16.00	98.92	3.8e+03	26	! ARX116767 Sequence 1890 from Patent	gb_pat:AR097017	+	16.00	98.43	4.0e+03	28	! AR097017 Sequence 80 from Patent
gb_pat:ARX117963	+	16.00	98.92	3.8e+03	26	! ARX117963 Sequence 3086 from Patent	gb_pat:AR125072	+	16.00	98.43	4.0e+03	28	! AR125072 Sequence 13 from Patent
gb_pat:ARX119405	+	16.00	98.92	3.8e+03	26	! ARX119405 Sequence 62 from Patent	gb_pat:AR150988	+	16.00	98.43	4.0e+03	28	! AR150988 Sequence 9 from Patent
gb_pat:ARX191907	+	16.00	98.92	3.8e+03	26	! ARX191907 Sequence 59 from Patent	gb_pat:AR156058	+	16.00	98.43	4.0e+03	28	! AR156058 Sequence 9 from Patent
gb_pat:ARX192142	+	16.00	98.92	3.8e+03	26	! ARX192142 Sequence 3 from Patent	gb_pat:AR172065	+	16.00	98.43	4.0e+03	28	! AR172065 Sequence 19 from Patent
gb_pat:ARX192143	+	16.00	98.92	3.8e+03	26	! ARX192143 Sequence 4 from Patent	gb_pat:AR173356	+	16.00	98.43	4.0e+03	28	! AR173356 Sequence 19 from Patent
gb_pat:ARX192146	+	16.00	98.92	3.8e+03	26	! ARX192146 Sequence 7 from Patent	gb_pat:ARX029525	+	16.00	98.43	4.0e+03	28	! ARX029525 Sequence 9 from Patent
gb_pat:ARX192149	+	16.00	98.92	3.8e+03	26	! ARX192149 Sequence 11 from Patent	gb_pat:ARX054734	+	16.00	98.43	4.0e+03	28	! ARX054734 Sequence 38 from Patent
gb_pat:ARX192150	+	16.00	98.92	3.8e+03	26	! ARX192150 Sequence 10 from Patent	gb_pat:ARX084098	+	16.00	98.43	4.0e+03	28	! ARX084098 Sequence 133 from Patent
gb_pat:ARX192152	+	16.00	98.92	3.8e+03	26	! ARX192152 Sequence 13 from Patent	gb_pat:ARX107611	+	16.00	98.43	4.0e+03	28	! ARX107611 Sequence 430 from Patent
gb_pat:ARX192153	+	16.00	98.92	3.8e+03	26	! ARX192153 Sequence 14 from Patent	gb_pat:ARX115674	+	16.00	98.43	4.0e+03	28	! ARX115674 Sequence 797 from Patent
gb_pat:ARX256080	-	16.00	98.92	3.8e+03	26	! ARX256080 Sequence 9 from Patent	gb_pat:ARX183800	+	16.00	98.43	4.0e+03	28	! ARX183800 Sequence 1553 from Patent
gb_pat:ARX300568	-	16.00	98.92	3.8e+03	26	! ARX300568 Sequence 74 from Patent	gb_pat:ARX252923	+	16.00	98.43	4.0e+03	28	! ARX252923 Sequence 393 from Patent
gb_pat:E28851	+	16.00	98.92	3.8e+03	26	! E28851 Process for preparing pri	gb_pat:ARX286499	+	16.00	98.43	4.0e+03	28	! ARX286499 Sequence 6 from Patent
gb_pat:E30823	-	16.00	98.92	3.8e+03	26	! E30823 Modified antibody Fab fra	gb_pat:ARX323027	+	16.00	98.43	4.0e+03	28	! ARX323027 Sequence 23 from Patent
gb_pat:E31574	-	16.00	98.92	3.8e+03	26	! E31574 Method for preparing DNA	gb_pat:ARX329101	+	16.00	98.43	4.0e+03	28	! ARX329101 Sequence 2 from Patent
gb_pat:E33560	-	16.00	98.92	3.8e+03	26	! E33560 Stress-responsive gene pr	gb_pat:ARX342689	+	16.00	98.43	4.0e+03	28	! ARX342689 Sequence 1 from Patent
gb_pat:E49558	-	16.00	98.92	3.8e+03	26	! E49558 Glucose dehydrogenase. 1/	gb_pat:IL06459	-	16.00	98.43	4.0e+03	28	! IL06459 Sequence 2 from Patent
gb_pat:E61349	+	16.00	98.92	3.8e+03	26	! E61349 Probe for detecting oligo	gb_pat:IL21978	-	16.00	98.43	4.0e+03	28	! IL21978 Sequence 64 from Patent
							gb_pat:IL21985	-	16.00	98.43	4.0e+03	28	! IL21985 Sequence 71 from Patent



gb_pat:187998	187998	4.0e+03	28	187998	Sequence 9 from patent US	gb_pat:AR138576	16.00	97.96	4.3e+03	30	AR138576	Sequence 101 from pa
gb_pat:A57879	57879	4.2e+03	29	A57879	Sequence 2 from Patent EP	gb_pat:AR138579	16.00	97.96	4.3e+03	30	AR138579	Sequence 104 from pa
gb_pat:A68962	68962	4.2e+03	29	A68962	Sequence 34 from Patent W	gb_pat:AR138591	16.00	97.96	4.3e+03	30	AR138591	Sequence 116 from pa
gb_pat:A69264	69264	4.2e+03	29	A69264	Sequence 3 from Patent WC	gb_pat:AR138594	16.00	97.96	4.3e+03	30	AR138594	Sequence 119 from pa
gb_pat:A71416	71416	4.2e+03	29	A71416	Sequence 27 from Patent W	gb_pat:AR138615	16.00	97.96	4.3e+03	30	AR138615	Sequence 140 from pa
gb_pat:A71941	71941	4.2e+03	29	A71941	Sequence 3 from Patent WC	gb_pat:AR138645	16.00	97.96	4.3e+03	30	AR138645	Sequence 170 from pa
gb_pat:A76894	76894	4.2e+03	29	A76894	Sequence 26 from Patent W	gb_pat:AR138654	16.00	97.96	4.3e+03	30	AR138654	Sequence 179 from pa
gb_pat:AR054999	54999	4.2e+03	29	AR054999	Sequence 37 from patent	gb_pat:AR138662	16.00	97.96	4.3e+03	30	AR138662	Sequence 187 from pa
gb_pat:AR095000	95000	4.2e+03	29	AR095000	Sequence 38 from patent	gb_pat:AR138674	16.00	97.96	4.3e+03	30	AR138674	Sequence 199 from pa
gb_pat:AR098149	98149	4.2e+03	29	AR098149	Sequence 25 from patent	gb_pat:AR138676	16.00	97.96	4.3e+03	30	AR138676	Sequence 201 from pa
gb_pat:AR098648	98648	4.2e+03	29	AR098648	Sequence 6 from patent	gb_pat:AX028510	16.00	97.96	4.3e+03	30	AX028510	Sequence 13 from Pat
gb_pat:AR101056	101056	4.2e+03	29	AR101056	Sequence 32 from patent	gb_pat:AX032823	16.00	97.96	4.3e+03	30	AX032823	Sequence 1 from Pat
gb_pat:AR101062	101062	4.2e+03	29	AR101062	Sequence 26 from patent	gb_pat:AX036047	16.00	97.96	4.3e+03	30	AX036047	Sequence 32 from Pat
gb_pat:AR101064	101064	4.2e+03	29	AR101064	Sequence 34 from patent	gb_pat:AX038497	16.00	97.96	4.3e+03	30	AX038497	Sequence 18 from Pat
gb_pat:AR116621	116621	4.2e+03	29	AR116621	Sequence 25 from patent	gb_pat:AX098446	16.00	97.96	4.3e+03	30	AX098446	Sequence 14 from Pat
gb_pat:AR131988	131988	4.2e+03	29	AR131988	Sequence 40 from patent	gb_pat:AX100424	16.00	97.96	4.3e+03	30	AX100424	Sequence 64 from Pat
gb_pat:AR147129	147129	4.2e+03	29	AR147129	Sequence 3 from patent	gb_pat:AX116662	16.00	97.96	4.3e+03	30	AX116662	Sequence 1785 from P
gb_pat:AR162080	162080	4.2e+03	29	AR162080	Sequence 8 from patent	gb_pat:AX117042	16.00	97.96	4.3e+03	30	AX117042	Sequence 2165 from P
gb_pat:AR166605	166605	4.2e+03	29	AR166605	Sequence 8 from patent	gb_pat:AX280450	16.00	97.96	4.3e+03	30	AX280450	Sequence 73 from Pat
gb_pat:AR176110	176110	4.2e+03	29	AR176110	Sequence 96 from patent	gb_pat:AX288069	16.00	97.96	4.3e+03	30	AX288069	Sequence 14 from Pat
gb_pat:AX011029	11029	4.2e+03	29	AX011029	Sequence 16 from Patent	gb_pat:AX288076	16.00	97.96	4.3e+03	30	AX288076	Sequence 21 from Pat
gb_pat:AX016303	16303	4.2e+03	29	AX016303	Sequence 6 from Patent	gb_pat:AX351705	16.00	97.96	4.3e+03	30	AX351705	Sequence 1 from Pat
gb_pat:AX048408	48408	4.2e+03	29	AX048408	Sequence 7 from Patent	gb_pat:AX351706	16.00	97.96	4.3e+03	30	AX351706	Sequence 2 from Pat
gb_pat:AX048409	48409	4.2e+03	29	AX048409	Sequence 8 from Patent	gb_pat:AX351707	16.00	97.96	4.3e+03	30	AX351707	Sequence 3 from Pat
gb_pat:AX052989	52989	4.2e+03	29	AX052989	Sequence 5 from Patent	gb_pat:AX351708	16.00	97.96	4.3e+03	30	AX351708	Sequence 4 from Pat
gb_pat:AX052994	52994	4.2e+03	29	AX052994	Sequence 8 from Patent	gb_pat:AX351710	16.00	97.96	4.3e+03	30	AX351710	Sequence 6 from Pat
gb_pat:AX078918	78918	4.2e+03	29	AX078918	Sequence 4 from Patent	gb_pat:AX351711	16.00	97.96	4.3e+03	30	AX351711	Sequence 7 from Pat
gb_pat:AX078919	78919	4.2e+03	29	AX078919	Sequence 5 from Patent	gb_pat:AX352051	16.00	97.96	4.3e+03	30	AX352051	Sequence 347 from Pa
gb_pat:AX268782	268782	4.2e+03	29	AX268782	Sequence 4 from Patent	gb_pat:BD008891	16.00	97.96	4.3e+03	30	BD008891	High level expressio
gb_pat:AX286501	286501	4.2e+03	29	AX286501	Sequence 8 from Patent	gb_pat:E05835	16.00	97.96	4.3e+03	30	E05835	DNA sequences of fragm
gb_pat:AX297747	297747	4.2e+03	29	AX297747	Sequence 9509 from Pat	gb_pat:E13629	16.00	97.96	4.3e+03	30	E13629	A part of polyribonuc
gb_pat:AX304639	304639	4.2e+03	29	AX304639	Sequence 153 from Pat	gb_pat:E39151	16.00	97.96	4.3e+03	30	E39151	Method for transformin
gb_pat:AX351955	351955	4.2e+03	29	AX351955	Sequence 251 from Pat	gb_pat:E59120	16.00	97.96	4.3e+03	30	E59120	Recombinant subunit va
gb_pat:AX353685	353685	4.2e+03	29	AX353685	Sequence 5 from Patent	gb_pat:I12382	16.00	97.96	4.3e+03	30	I12382	Sequence 5 from patent
gb_pat:BD006075	606075	4.2e+03	29	BD006075	Polynucleotide vaccine	gb_pat:I14394	16.00	97.96	4.3e+03	30	I14394	Sequence 18 from patent
gb_pat:BD009817	9817	4.2e+03	29	BD009817	Avian polynucleotide va	gb_pat:I38507	16.00	97.96	4.3e+03	30	I38507	Sequence 85 from patent
gb_pat:I65795	65795	4.2e+03	29	I65795	Sequence 13 from patent U	gb_pat:I39988	16.00	97.96	4.3e+03	30	I39988	Sequence 41 from patent
gb_pat:A36778	36778	4.3e+03	30	A36778	Oligonucleotide GP67-SC.1	gb_pat:I49601	16.00	97.96	4.3e+03	30	I49601	Sequence 43 from patent
gb_pat:A38144	38144	4.3e+03	30	A38144	Sequence 6 from Patent EP	gb_pat:I56982	16.00	97.96	4.3e+03	30	I56982	Sequence 85 from patent
gb_pat:A91182	91182	4.3e+03	30	A91182	Sequence 16 from Patent W	gb_pat:I59848	16.00	97.96	4.3e+03	30	I59848	Sequence 85 from patent
gb_pat:A83875	83875	4.3e+03	30	A83875	Sequence 10 from Patent W	gb_pat:I75175	16.00	97.96	4.3e+03	30	I75175	Sequence 85 from patent
gb_pat:A93376	93376	4.3e+03	30	A93376	Sequence 11 from Patent W	gb_pat:I91906	16.00	97.96	4.3e+03	30	I91906	Sequence 4 from patent
gb_pat:A93586	93586	4.3e+03	30	A93586	Sequence 16 from Patent W	gb_pat:MHSD06M	16.00	97.96	4.3e+03	30	MHSD06M	H.sapiens mitochondria
gb_pat:AR000059	000059	4.3e+03	30	AR000059	Sequence 7 from patent	gb_pat:A04979	16.00	97.74	4.4e+03	31	A04979	Artificial sequence fo
gb_pat:AR016018	16018	4.3e+03	30	AR016018	Sequence 50 from patent	gb_pat:A04980	16.00	97.74	4.4e+03	31	A04980	Artificial sequence fo
gb_pat:AR016852	16852	4.3e+03	30	AR016852	Sequence 85 from patent	gb_pat:A08914	16.00	97.74	4.4e+03	31	A08914	H.sapiens (haplotype 3
gb_pat:AR020878	20878	4.3e+03	30	AR020878	Sequence 28 from patent	gb_pat:A74270	16.00	97.74	4.4e+03	31	A74270	Sequence 26 from Paten
gb_pat:AR023796	23796	4.3e+03	30	AR023796	Sequence 85 from patent	gb_pat:AR028224	16.00	97.74	4.4e+03	31	AR028224	Sequence 73 from pat
gb_pat:AR027201	27201	4.3e+03	30	AR027201	Sequence 85 from patent	gb_pat:AR029622	16.00	97.74	4.4e+03	31	AR029622	Sequence 7 from pat
gb_pat:AR028168	28168	4.3e+03	30	AR028168	Sequence 17 from patent	gb_pat:AR063949	16.00	97.74	4.4e+03	31	AR063949	Sequence 48 from pat
gb_pat:AR028174	28174	4.3e+03	30	AR028174	Sequence 23 from patent	gb_pat:AR079777	16.00	97.74	4.4e+03	31	AR079777	Sequence 101 from pa
gb_pat:AR028176	28176	4.3e+03	30	AR028176	Sequence 25 from patent	gb_pat:AR081307	16.00	97.74	4.4e+03	31	AR081307	Sequence 101 from pa
gb_pat:AR028188	28188	4.3e+03	30	AR028188	Sequence 37 from patent	gb_pat:AR091411	16.00	97.74	4.4e+03	31	AR091411	Sequence 1 from pat
gb_pat:AR028191	28191	4.3e+03	30	AR028191	Sequence 40 from patent	gb_pat:AR125616	16.00	97.74	4.4e+03	31	AR125616	Sequence 1 from pat
gb_pat:AR028212	28212	4.3e+03	30	AR028212	Sequence 91 from patent	gb_pat:AR138627	16.00	97.74	4.4e+03	31	AR138627	Sequence 152 from pa
gb_pat:AR028242	28242	4.3e+03	30	AR028242	Sequence 61 from patent	gb_pat:AR149549	16.00	97.74	4.4e+03	31	AR149549	Sequence 101 from pa
gb_pat:AR028251	28251	4.3e+03	30	AR028251	Sequence 100 from Paten	gb_pat:AR170667	16.00	97.74	4.4e+03	31	AR170667	Sequence 101 from pa
gb_pat:AR028259	28259	4.3e+03	30	AR028259	Sequence 128 from patent	gb_pat:AR195657	16.00	97.74	4.4e+03	31	AR195657	Sequence 5 from Pat
gb_pat:AR028271	28271	4.3e+03	30	AR028271	Sequence 108 from patent	gb_pat:AX069293	16.00	97.74	4.4e+03	31	AX069293	Sequence 6 from Pat
gb_pat:AR028273	28273	4.3e+03	30	AR028273	Sequence 122 from patent	gb_pat:AX100366	16.00	97.74	4.4e+03	31	AX100366	Sequence 5 from Pat
gb_pat:AR038488	38488	4.3e+03	30	AR038488	Sequence 85 from patent	gb_pat:AX113882	16.00	97.74	4.4e+03	31	AX113882	Sequence 303 from Pa
gb_pat:AR064630	64630	4.3e+03	30	AR064630	Sequence 85 from patent	gb_pat:AX115915	16.00	97.74	4.4e+03	31	AX115915	Sequence 1038 from Pa
gb_pat:AR064882	64882	4.3e+03	30	AR064882	Sequence 7 from patent	gb_pat:AX159660	16.00	97.74	4.4e+03	31	AX159660	Sequence 203 from Pa
gb_pat:AR064925	64925	4.3e+03	30	AR064925	Sequence 50 from patent	gb_pat:AX196238	16.00	97.74	4.4e+03	31	AX196238	Sequence 69 from Pat
gb_pat:AR067555	67555	4.3e+03	30	AR067555	Sequence 85 from patent	gb_pat:AX221242	16.00	97.74	4.4e+03	31	AX221242	Sequence 6684 from Pa
gb_pat:AR082883	82883	4.3e+03	30	AR082883	Sequence 16 from patent	gb_pat:AX223359	16.00	97.74	4.4e+03	31	AX223359	Sequence 8801 from Pa
gb_pat:AR084895	84895	4.3e+03	30	AR084895	Sequence 40 from patent	gb_pat:AX223469	16.00	97.74	4.4e+03	31	AX223469	Sequence 8911 from Pa
gb_pat:AR103197	3197	4.3e+03	30	AR103197	Sequence 91 from patent	gb_pat:AX223500	16.00	97.74	4.4e+03	31	AX223500	Sequence 8942 from Pa
gb_pat:AR110027	110027	4.3e+03	30	AR110027	Sequence 28 from patent	gb_pat:AX229041	16.00	97.74	4.4e+03	31	AX229041	Sequence 2413 from Pa
gb_pat:AR134337	134337	4.3e+03	30	AR134337	Sequence 20 from patent	gb_pat:AX248129	16.00	97.74	4.4e+03	31	AX248129	Sequence 208 from Pa
gb_pat:AR138571	138571	4.3e+03	30	AR138571	Sequence 96 from patent U	gb_pat:AX248178	16.00	97.74	4.4e+03	31	AX248178	Sequence 257 from Pa



gb_pat:BD002609	+	16.00	97.74	4.4e+03	31	BD002609	Gene composition and me	gb_pat:AX007104	+	16.00	97.32	4.7e+03	33	AX007104	Sequence 4 from Pat
gb_pat:BD002609	-	16.00	97.74	4.4e+03	31	BD002609	Gene composition and me	gb_pat:AX009245	-	16.00	97.32	4.7e+03	33	AX009245	Sequence 1 from Pat
gb_pat:BD002745	-	16.00	97.74	4.4e+03	31	BD002745	Gene composition and me	gb_pat:AX0040645	-	16.00	97.32	4.7e+03	33	AX0040645	Sequence 2 from Pat
gb_pat:E05836	+	16.00	97.74	4.4e+03	31	E05836	DNA sequence of fragment	gb_pat:AX076479	-	16.00	97.32	4.7e+03	33	AX076479	Sequence 5 from Pat
gb_pat:I41011	+	16.00	97.74	4.4e+03	31	I41011	Sequence 7 from patent US	gb_pat:AX080517	-	16.00	97.32	4.7e+03	33	AX080517	Sequence 5 from Pat
gb_pat:I92782	+	16.00	97.74	4.4e+03	31	I92782	Sequence 48 from patent U	gb_pat:AX288171	+	16.00	97.32	4.7e+03	33	AX288171	Sequence 42 from Pat
gb_pat:AX48512	+	16.00	97.53	4.5e+03	32	AX48512	Sequence 14 from Patent W	gb_pat:BD006668	+	16.00	97.32	4.7e+03	33	BD006668	Transporter gene OAT
gb_pat:AX76289	+	16.00	97.53	4.5e+03	32	AX76289	Sequence 12 from Patent W	gb_pat:BD008459	+	16.00	97.32	4.7e+03	33	BD008459	Tartrigester adenovirus
gb_pat:AX9130	-	16.00	97.53	4.5e+03	32	AX9130	Sequence 10 from Patent W	gb_pat:BD010714	-	16.00	97.32	4.7e+03	33	BD010714	Helicobacter polypep
gb_pat:AR003426	-	16.00	97.53	4.5e+03	32	AR003426	Sequence 61 from patent	gb_pat:E59628	-	16.00	97.32	4.7e+03	33	E59628	Cloning vector for mRN
gb_pat:AR007545	+	16.00	97.53	4.5e+03	32	AR007545	Sequence 18 from patent	gb_pat:II2575	+	16.00	97.32	4.7e+03	33	II2575	Clonence 21 from patent
gb_pat:AR012205	+	16.00	97.53	4.5e+03	32	AR012205	Sequence 18 from patent	gb_pat:I38515	+	16.00	97.32	4.7e+03	33	I38515	Sequence 93 from patent
gb_pat:AR014032	+	16.00	97.53	4.5e+03	32	AR014032	Sequence 68 from patent	gb_pat:I38517	+	16.00	97.32	4.7e+03	33	I38517	Sequence 95 from patent
gb_pat:AR020449	+	16.00	97.53	4.5e+03	32	AR020449	Sequence 1 from patent	gb_pat:I45569	-	16.00	97.32	4.7e+03	33	I45569	Sequence 4 from patent
gb_pat:AR022620	+	16.00	97.53	4.5e+03	32	AR022620	Sequence 18 from patent	gb_pat:I56990	+	16.00	97.32	4.7e+03	33	I56990	Sequence 93 from patent
gb_pat:AR023893	+	16.00	97.53	4.5e+03	32	AR023893	Sequence 18 from patent	gb_pat:I56992	+	16.00	97.32	4.7e+03	33	I56992	Sequence 95 from patent
gb_pat:AR029811	+	16.00	97.53	4.5e+03	32	AR029811	Sequence 48 from patent	gb_pat:I59856	+	16.00	97.32	4.7e+03	33	I59856	Sequence 93 from patent
gb_pat:AR038820	-	16.00	97.53	4.5e+03	32	AR038820	Sequence 9 from patent	gb_pat:I59858	+	16.00	97.32	4.7e+03	33	I59858	Sequence 95 from patent
gb_pat:AR051171	-	16.00	97.53	4.5e+03	32	AR051171	Sequence 52 from patent	gb_pat:I72106	+	16.00	97.32	4.7e+03	33	I72106	Sequence 91 from patent
gb_pat:AR051291	+	16.00	97.53	4.5e+03	32	AR051291	Sequence 8 from patent	gb_pat:I75183	+	16.00	97.32	4.7e+03	33	I75183	Sequence 21 from patent
gb_pat:AR059059	+	16.00	97.53	4.5e+03	32	AR059059	Sequence 18 from patent	gb_pat:I75185	+	16.00	97.32	4.7e+03	33	I75185	Sequence 95 from patent
gb_pat:AR093927	+	16.00	97.53	4.5e+03	32	AR093927	Sequence 31 from patent	gb_pat:I96158	+	16.00	97.32	4.7e+03	33	I96158	Sequence 8 from patent
gb_pat:AR094202	-	16.00	97.53	4.5e+03	32	AR094202	Sequence 17 from patent	gb_pat:AX21870	-	16.00	97.12	4.8e+03	34	AX21870	oligonucleotide primer
gb_pat:AR123494	-	16.00	97.53	4.5e+03	32	AR123494	Sequence 3 from patent	gb_pat:AX3805	-	16.00	97.12	4.8e+03	34	AX3805	Artificial DNA for pri
gb_pat:AR138577	-	16.00	97.53	4.5e+03	32	AR138577	Sequence 102 from patent	gb_pat:AX40132	+	16.00	97.12	4.8e+03	34	AX40132	Sequence 8 from Patent
gb_pat:AR166732	+	16.00	97.53	4.5e+03	32	AR166732	Sequence 18 from patent	gb_pat:AX48920	+	16.00	97.12	4.8e+03	34	AX48920	Sequence 11 from Patent
gb_pat:AX008524	+	16.00	97.53	4.5e+03	32	AX008524	Sequence 176 from Patent	gb_pat:AX57100	-	16.00	97.12	4.8e+03	34	AX57100	Sequence 4 from Patent
gb_pat:AX104299	-	16.00	97.53	4.5e+03	32	AX104299	Sequence 491 from Patent	gb_pat:AX59838	-	16.00	97.12	4.8e+03	34	AX59838	Sequence 8 from Patent
gb_pat:AX120020	-	16.00	97.53	4.5e+03	32	AX120020	Sequence 3 from Patent	gb_pat:AX60180	-	16.00	97.12	4.8e+03	34	AX60180	Sequence 58 from Patent
gb_pat:AX147210	+	16.00	97.53	4.5e+03	32	AX147210	Sequence 10 from Patent	gb_pat:AX62749	+	16.00	97.12	4.8e+03	34	AX62749	Sequence 13 from Patent
gb_pat:AX188790	+	16.00	97.53	4.5e+03	32	AX188790	Sequence 1 from Patent	gb_pat:AX63578	+	16.00	97.12	4.8e+03	34	AX63578	Sequence 19 from Patent
gb_pat:AX196220	+	16.00	97.53	4.5e+03	32	AX196220	Sequence 51 from Patent	gb_pat:AX67238	-	16.00	97.12	4.8e+03	34	AX67238	Sequence 22 from Patent
gb_pat:AX236574	+	16.00	97.53	4.5e+03	32	AX236574	Sequence 267 from Patent	gb_pat:AX79530	-	16.00	97.12	4.8e+03	34	AX79530	Sequence 68 from Patent
gb_pat:AX259612	+	16.00	97.53	4.5e+03	32	AX259612	Sequence 32 from Patent	gb_pat:AR027109	-	16.00	97.12	4.8e+03	34	AR027109	Sequence 9 from patent
gb_pat:AX268779	-	16.00	97.53	4.5e+03	32	AX268779	Sequence 1 from Patent	gb_pat:AR047929	-	16.00	97.12	4.8e+03	34	AR047929	Sequence 11 from pat
gb_pat:AX355459	-	16.00	97.53	4.5e+03	32	AX355459	Sequence 487 from Patent	gb_pat:AR053489	+	16.00	97.12	4.8e+03	34	AR053489	Sequence 80 from pat
gb_pat:E03507	-	16.00	97.53	4.5e+03	32	E03507	Primer for amplifying chr	gb_pat:AR094754	+	16.00	97.12	4.8e+03	34	AR094754	Sequence 68 from pat
gb_pat:E13138	+	16.00	97.53	4.5e+03	32	E13138	PCR primer. 6/1998	gb_pat:AR126758	+	16.00	97.12	4.8e+03	34	AR126758	Sequence 1 from pat
gb_pat:I06269	+	16.00	97.53	4.5e+03	32	I06269	Sequence 27 from Patent E	gb_pat:AR127732	+	16.00	97.12	4.8e+03	34	AR127732	Sequence 150 from pat
gb_pat:I16939	-	16.00	97.53	4.5e+03	32	I16939	Sequence 8 from patent US	gb_pat:AR162956	-	16.00	97.12	4.8e+03	34	AR162956	Sequence 11 from pat
gb_pat:I121215	-	16.00	97.53	4.5e+03	32	I121215	Sequence 61 from patent US	gb_pat:AX001019	-	16.00	97.12	4.8e+03	34	AX001019	Sequence 1 from Pat
gb_pat:I121848	-	16.00	97.53	4.5e+03	32	I121848	Sequence 3 from patent US	gb_pat:AX068844	-	16.00	97.12	4.8e+03	34	AX068844	Sequence 4 from Pat
gb_pat:I121982	-	16.00	97.53	4.5e+03	32	I121982	Sequence 68 from patent U	gb_pat:AX100352	-	16.00	97.12	4.8e+03	34	AX100352	Sequence 35 from Pat
gb_pat:I145733	+	16.00	97.53	4.5e+03	32	I145733	Sequence 8 from patent US	gb_pat:AX180829	+	16.00	97.12	4.8e+03	34	AX180829	Sequence 6 from Pat
gb_pat:I51731	-	16.00	97.53	4.5e+03	32	I51731	Sequence 52 from patent U	gb_pat:AX183973	+	16.00	97.12	4.8e+03	34	AX183973	Sequence 1726 from P
gb_pat:I60422	-	16.00	97.53	4.5e+03	32	I60422	Sequence 9 from patent US	gb_pat:AX202205	+	16.00	97.12	4.8e+03	34	AX202205	Sequence 10 from Pat
gb_pat:I74482	-	16.00	97.53	4.5e+03	32	I74482	Sequence 61 from patent U	gb_pat:AX350106	+	16.00	97.12	4.8e+03	34	AX350106	Sequence 629 from Pa
gb_pat:I75221	+	16.00	97.53	4.5e+03	32	I75221	Sequence 18 from patent U	gb_pat:AX351958	+	16.00	97.12	4.8e+03	34	AX351958	Sequence 254 from Pa
gb_pat:I75246	+	16.00	97.53	4.5e+03	32	I75246	Sequence 18 from patent U	gb_pat:AX351961	+	16.00	97.12	4.8e+03	34	AX351961	Sequence 257 from Pa
gb_pat:I88820	+	16.00	97.32	4.7e+03	33	I88820	Sequence 28 from patent U	gb_pat:AX351962	+	16.00	97.12	4.8e+03	34	AX351962	Sequence 258 from Pa
gb_pat:AX30241	+	16.00	97.32	4.7e+03	33	AX30241	H. contortus 40 kb gene PC	gb_pat:AX351963	+	16.00	97.12	4.8e+03	34	AX351963	Sequence 259 from Pa
gb_pat:AX68665	-	16.00	97.32	4.7e+03	33	AX68665	Sequence 11 from Patent W	gb_pat:AX351964	+	16.00	97.12	4.8e+03	34	AX351964	Sequence 260 from Pa
gb_pat:AX68666	-	16.00	97.32	4.7e+03	33	AX68666	Sequence 12 from Patent W	gb_pat:AX351966	-	16.00	97.12	4.8e+03	34	AX351966	Sequence 262 from Pa
gb_pat:AR016860	+	16.00	97.32	4.7e+03	33	AR016860	Sequence 93 from patent	gb_pat:AX351967	+	16.00	97.12	4.8e+03	34	AX351967	Sequence 263 from Pa
gb_pat:AR016862	+	16.00	97.32	4.7e+03	33	AR016862	Sequence 95 from patent	gb_pat:AX351968	+	16.00	97.12	4.8e+03	34	AX351968	Sequence 264 from Pa
gb_pat:AR020886	+	16.00	97.32	4.7e+03	33	AR020886	Sequence 93 from patent	gb_pat:BD009365	+	16.00	97.12	4.8e+03	34	BD009365	Method for preparing Fab
gb_pat:AR020888	+	16.00	97.32	4.7e+03	33	AR020888	Sequence 95 from patent	gb_pat:AX30822	+	16.00	97.12	4.8e+03	34	AX30822	Modified antibody Fab
gb_pat:AR027209	+	16.00	97.32	4.7e+03	33	AR027209	Sequence 93 from patent	gb_pat:E59675	+	16.00	97.12	4.8e+03	34	E59675	Method for differentia
gb_pat:AR027211	+	16.00	97.32	4.7e+03	33	AR027211	Sequence 95 from patent	gb_pat:E59843	-	16.00	97.12	4.8e+03	34	E59843	Method for imparting h
gb_pat:AR035267	+	16.00	97.32	4.7e+03	33	AR035267	Sequence 8 from patent	gb_pat:I25955	+	16.00	97.12	4.8e+03	34	I25955	Sequence 8 from patent
gb_pat:AR036358	+	16.00	97.32	4.7e+03	33	AR036358	Sequence 21 from patent	gb_pat:I72227	+	16.00	97.12	4.8e+03	34	I72227	Sequence 11 from patent
gb_pat:AR038496	+	16.00	97.32	4.7e+03	33	AR038496	Sequence 93 from patent	gb_pat:I75401	+	16.00	97.12	4.8e+03	34	I75401	Sequence 150 from pate
gb_pat:AR038498	+	16.00	97.32	4.7e+03	33	AR038498	Sequence 95 from patent	gb_sts:C75781	+	16.00	97.12	4.8e+03	34	C75781	Homo sapiens STS WI-11
gb_pat:AR064060	+	16.00	97.32	4.7e+03	33	AR064060	Sequence 18 from patent	gb_pat:AX24382	-	16.00	96.93	4.9e+03	35	AX24382	dtl17 adaptor primer. 3
gb_pat:AR064638	+	16.00	97.32	4.7e+03	33	AR064638	Sequence 93 from patent	gb_pat:AX27645	-	16.00	96.93	4.9e+03	35	AX27645	Hybrid dtl17-adaptor pr
gb_pat:AR064640	+	16.00	97.32	4.7e+03	33	AR064640	Sequence 95 from patent	gb_pat:AX37244	-	16.00	96.93	4.9e+03	35	AX37244	Sequence 13 from Patent
gb_pat:AR067563	+	16.00	97.32	4.7e+03	33	AR067563	Sequence 93 from patent	gb_pat:AX40373	-	16.00	96.93	4.9e+03	35	AX40373	Sequence 29 from Patent
gb_pat:AR067565	+	16.00	97.32	4.7e+03	33	AR067565	Sequence 95 from patent	gb_pat:AX42335	-	16.00	96.93	4.9e+03	35	AX42335	Sequence 7 from Patent
gb_pat:AR095931	+	16.00	97.32	4.7e+03	33	AR095931	Sequence 4 from patent	gb_pat:AX42384	-	16.00	96.93	4.9e+03	35	AX42384	Sequence 7 from Patent
gb_pat:AR118072	-	16.00	97.32	4.7e+03	33	AR118072	Sequence 14 from patent	gb_pat:AX46467	-	16.00	96.93	4.9e+03	35	AX46467	Sequence 5 from Patent
gb_pat:AR118077	+	16.00	97.32	4.7e+03	33	AR118077	Sequence 8 from patent	gb_pat:AX62983	+	16.00	96.93	4.9e+03	35	AX62983	Sequence 2 from Patent
gb_pat:AR172404	+	16.00	97.32	4.7e+03	33	AR172404	Sequence 1 from patent	gb_pat:AX67228	-	16.00	96.93	4.9e+03	35	AX67228	Sequence 12 from Patent
gb_pat:AX002722	-	16.00	97.32	4.7e+03	33	AX002722	Sequence 5 from Patent	gb_pat:AX68608	-	16.00	96.93	4.9e+03	35	AX68608	Sequence 8 from Patent



gb\_pat:A75916 - 16.00 96.93 4.9e+03 35 ! A75916 Sequence 16 from Patent W  
 gb\_pat:A76306 - 16.00 96.93 4.9e+03 35 ! A76306 Sequence 12 from Patent W  
 gb\_pat:A76310 - 16.00 96.93 4.9e+03 35 ! A76310 Sequence 16 from Patent W  
 gb\_pat:AR012367 - 16.00 96.93 4.9e+03 35 ! AR012367 Sequence 5 from patent  
 gb\_pat:AR023670 - 16.00 96.93 4.9e+03 35 ! AR023670 Sequence 68 from patent  
 gb\_pat:AR029830 - 16.00 96.93 4.9e+03 35 ! AR029830 Sequence 19 from patent  
 gb\_pat:AR037793 - 16.00 96.93 4.9e+03 35 ! AR037793 Sequence 11 from patent  
 gb\_pat:AR038057 - 16.00 96.93 4.9e+03 35 ! AR038057 Sequence 7 from patent  
 gb\_pat:AR040751 - 16.00 96.93 4.9e+03 35 ! AR040751 Sequence 28 from patent  
 gb\_pat:AR043620 - 16.00 96.93 4.9e+03 35 ! AR043620 Sequence 5 from patent  
 em\_pat:E10418 - 16.00 96.93 4.9e+03 35 ! E10418 Primer. 9/2000

seq\_name: gb\_pat:AR096851

seq\_documentation\_block:  
 LOCUS AR096851 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 2 from patent US 6008400.

ACCESSION AR096851

VERSION AR096851.1 GI:10026020

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 2 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g -1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR096851 ..

Align seg 1/1 to: AR096851 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAAAAAAAAAT 10

seq\_name: gb\_pat:AR096852

seq\_documentation\_block:

LOCUS AR096852 10 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 3 from patent US 6008400.

ACCESSION AR096852

VERSION AR096852.1 GI:10026022

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Scaringe,S. and Caruthers,M.H.

TITLE Orthoester reagents for use as protecting groups in oligonucleotide

JOURNAL synthesis

FEATURES Patent: US 6008400-A 3 28-DEC-1999;

source Location/Qualifiers

BASE COUNT 1. .10

ORIGIN /organism="unknown"

alignment\_scores: 9 a 0 c 0 g 1 t

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0



BASE COUNT 1 a 0 c 1 g 8 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AX301521/rev ..

Align seg 1/1 to reverse of: AX301521 from: 1 to: 10

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: gb\_pat:AX061320

seq\_documentation\_block:  
LOCUS AX061320 11 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 54 from Patent W00100816.  
ACCESSION AX061320  
VERSION AX061320.1 GI:12406455

KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
artificial sequence.

REFERENCE 1 (bases 1 to 11)  
AUTHORS Lexow,P.  
TITLE Methods of cloning and producing fragment chains with readable

information content  
JOURNAL Patent: WO 0100816-A 54 04-JAN-2001;  
Complete Genomics AS (NO)

FEATURES  
source Location/Qualifiers  
1..11  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

BASE COUNT 10 a 1 c 0 g 0 t  
ORIGIN  
/note="1 Starting fragment, position 2"

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AX061320 ..

Align seg 1/1 to: AX061320 from: 1 to: 11

1 LysLysAsn 3  
|||||  
3 AAAAAAAC 11

seq\_name: gb\_pat:AX100363

seq\_documentation\_block:  
LOCUS AX100363 11 bp DNA linear PAT 10-APR-2001  
DEFINITION Sequence 3 from Patent W00121833.

ACCESSION AX100363  
VERSION AX100363.1 GI:13619379

KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
artificial sequence.

REFERENCE 1 (bases 1 to 11)  
AUTHORS Bayliffe,A.I., Docter,E., Kelly,S.J. and Robertson,N.H.

TITLE Assay for detection of human cfr allele variants using specific  
diagnostic primers

JOURNAL Patent: WO 0121833-A 3 29-MAR-2001;  
Astrazeneca AB (SE)

FEATURES  
source Location/Qualifiers  
1..11  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

BASE COUNT 8 a 1 c 0 g 1 t 1 others  
ORIGIN  
/note="PCR primer"

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AX100363 ..

Align seg 1/1 to: AX100363 from: 1 to: 11

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: gb\_pat:AS8733

seq\_documentation\_block:

LOCUS AS8733 12 bp DNA linear PAT 06-MAR-1998  
DEFINITION Sequence 13 from Patent W09641812.

ACCESSION AS8733

VERSION AS8733.1 GI:3714275

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 12)

AUTHORS Brownlee,G.G., Lowe,G. and Fodor,E.

TITLE OLIGONUCLEOTIDE PHOSPHORYLATION METHOD AND PRODUCTS

JOURNAL Patent: WO 9641812-A 13 27-DEC-1996;

ISIS INNOVATION (GB)

FEATURES  
source Location/Qualifiers  
1..12  
/organism="unidentified"

modified\_base 1..2  
/db\_xref="taxon:32644"

BASE COUNT 0 a 0 c 1 g 11 t

ORIGIN  
/mod\_base=OTHER

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AS8733/rev ..

Align seg 1/1 to reverse of: AS8733 from: 1 to: 12

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: gb\_pat:AR162081

seq\_documentation\_block:

LOCUS AR162081 12 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 9 from patent US 6258558.

ACCESSION AR162081

VERSION AR162081.1 GI:16229145

KEYWORDS







```

KEYWORDS      Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 13)
AUTHORS        Chee,M., Gingers,T.R., Fodor,S.P.A., Hubble,E.A. and Morris,M.S.
TITLE          Array of nucleic acid probes on biological chips for diagnosis of
                HIV and methods of using the same
JOURNAL        Patent: US 5861242-A 19 19-JAN-1999;
FEATURES       Location/Qualifiers
                source
                1..13
                /organism="unknown"
BASE COUNT     1 a      1 c      1 g      10 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x AR029782/rev ..
Align seg 1/1 to reverse of: AR029782 from: 1 to: 13
1 LysLysAsn 3
|||||
13 AAAAAAAC 5
seq_name: gb_pat:AR175284
seq_documentation_block:
LOCUS      AR175284      13 bp      DNA      linear      PAT 17-DEC-2001
DEFINITION Sequence 7 from patent US 6309823.
ACCESSION  AR175284
VERSION     AR175284.1 GI:17916583
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 13)
AUTHORS      Cronin,M.T., Miyada,C.G., Hubbell,E.A., Chee,M., Fodor,S.P.A.,
                Huang,X.C., Lipshutz,R.J., Lobban,P.E., Morris,M.S. and
                Sheldon,E.L.
TITLE        Arrays of nucleic acid probes for analyzing biotransformation genes
                and methods of using the same
JOURNAL      Patent: US 6309823-A 7 30-OCT-2001;
FEATURES     Location/Qualifiers
                source
                1..13
                /organism="unknown"
BASE COUNT   1 a      1 c      1 g      10 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x AR175284/rev ..
Align seg 1/1 to reverse of: AR175284 from: 1 to: 13
1 LysLysAsn 3
|||||
13 AAAAAAAC 5
seq_name: gb_pat:AX061332
seq_documentation_block:
LOCUS      AX061332      13 bp      DNA      linear      PAT 22-JAN-2001
DEFINITION Sequence 7 from patent US 6309823.
ACCESSION  AX061332
VERSION     AX061332.1 GI:12406467
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE    1 (bases 1 to 13)
AUTHORS      Chee,M., Gingers,T.R., Fodor,S.P.A., Hubble,E.A. and Morris,M.S.
TITLE        Array of nucleic acid probes on biological chips for diagnosis of
                HIV and methods of using the same
JOURNAL      Patent: US 5861242-A 19 19-JAN-1999;
FEATURES     Location/Qualifiers
                source
                1..13
                /organism="unknown"
BASE COUNT   1 a      1 c      1 g      10 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x AX061332/rev ..
Align seg 1/1 to reverse of: AX061332 from: 1 to: 13
1 LysLysAsn 3
|||||
9 AAAAAAAC 1
seq_name: gb_pat:Al2784
seq_documentation_block:
LOCUS      Al2784      14 bp      DNA      linear      PAT 28-APR-1994
DEFINITION Oligonucleotide from clone PHD 116.
ACCESSION  Al2784
VERSION     Al2784.1 GI:512048
KEYWORDS
SOURCE      synthetic construct.
ORGANISM    synthetic construct.
REFERENCE    1 (bases 1 to 14)
AUTHORS
TITLE        A DNA SEQUENCE
JOURNAL      Patent: WO 8605804-A 15 09-OCT-1986;
FEATURES     Location/Qualifiers
                source
                1..14
                /organism="synthetic construct"
                /db_xref="taxon:32630"
BASE COUNT   3 a      4 c      0 g      7 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x Al2784/rev ..
Align seg 1/1 to reverse of: Al2784 from: 1 to: 14
1 LysLysAsn 3
|||||
10 AAGAAGAAT 2

```

```

DEFINITION      Sequence 66 from Patent WO0100816.
ACCESSION       AX061332
VERSION         AX061332.1 GI:12406467
KEYWORDS
SOURCE          synthetic construct.
ORGANISM        synthetic construct.
REFERENCE        1 (bases 1 to 13)
AUTHORS         Lexow,P.
TITLE           Methods of cloning and producing fragment chains with readable
                information content
JOURNAL         Patent: WO 0100816-A 66 04-JAN-2001;
FEATURES        Complete Genomics AS (NO)
                Location/Qualifiers
                source
                1..13
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="Fragment1, position 8.2"
BASE COUNT      0 a      1 c      2 g      10 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x AX061332/rev ..
Align seg 1/1 to reverse of: AX061332 from: 1 to: 13
1 LysLysAsn 3
|||||
9 AAAAAAAC 1
seq_name: gb_pat:Al2784
seq_documentation_block:
LOCUS      Al2784      14 bp      DNA      linear      PAT 28-APR-1994
DEFINITION Oligonucleotide from clone PHD 116.
ACCESSION  Al2784
VERSION     Al2784.1 GI:512048
KEYWORDS
SOURCE      synthetic construct.
ORGANISM    synthetic construct.
REFERENCE    1 (bases 1 to 14)
AUTHORS
TITLE        A DNA SEQUENCE
JOURNAL      Patent: WO 8605804-A 15 09-OCT-1986;
FEATURES     Location/Qualifiers
                source
                1..14
                /organism="synthetic construct"
                /db_xref="taxon:32630"
BASE COUNT   3 a      4 c      0 g      7 t
ORIGIN
alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333       Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x Al2784/rev ..
Align seg 1/1 to reverse of: Al2784 from: 1 to: 14
1 LysLysAsn 3
|||||
10 AAGAAGAAT 2

```



```
seq_name: gb_pat:A88063
seq_documentation_block:
LOCUS A88063 14 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 211 from Patent WO9833904.
ACCESSION A88063
VERSION A88063.1 GI:6736633
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 211 06-AUG-1998;
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source
1..14
/organism="unidentified"
/db_xref="taxon:32644"
1 t
BASE COUNT 10 a 2 c 1 g 1 t
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x A88063 ..
Align seg 1/1 to: A88063 from: 1 to: 14

1 LysLysAsn 3
|||||
2 AAAAAAAC 10

seq_name: gb_pat:A88065
seq_documentation_block:
LOCUS A88065 14 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 213 from Patent WO9833904.
ACCESSION A88065
VERSION A88065.1 GI:6736635
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 213 06-AUG-1998;
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source
1..14
/organism="unidentified"
/db_xref="taxon:32644"
1 t
BASE COUNT 10 a 2 c 1 g 1 t
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x A88065 ..
Align seg 1/1 to: A88065 from: 1 to: 14

1 LysLysAsn 3
|||||
2 AAAAAAAC 10

seq_name: gb_pat:A88107
seq_documentation_block:
LOCUS A88107 14 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 255 from Patent WO9833904.
ACCESSION A88107
VERSION A88107.1 GI:6736677
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 255 06-AUG-1998;
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source
1..14
/organism="unidentified"
/db_xref="taxon:32644"
1 t
BASE COUNT 10 a 2 c 1 g 1 t
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x A88107 ..
Align seg 1/1 to: A88107 from: 1 to: 14

1 LysLysAsn 3
|||||
2 AAAAAAAC 10

seq_name: gb_pat:A88110
seq_documentation_block:
LOCUS A88110 14 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 258 from Patent WO9833904.
ACCESSION A88110
VERSION A88110.1 GI:6736680
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 258 06-AUG-1998;
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source
1..14
/organism="unidentified"
/db_xref="taxon:32644"
2 t
BASE COUNT 9 a 2 c 1 g 2 t
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x A88110 ..
```



Align seg 1/1 to: A88110 from: 1 to: 14

1 LysLysAsn 3  
|||||  
6 AAAAAAAC 14

seq\_name: gb\_pat:A90030

seq\_documentation\_block:

LOCUS A90030 14 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 211 from Patent EP0856579.  
ACCESSION A90030  
VERSION A90030.1 GI:6738544

KEYWORDS  
SOURCE  
ORGANISM  
unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 211 05-AUG-1998;  
BIODOSTIK GES (DE)

FEATURES  
source Location/Qualifiers  
1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 10 a 2 c 1 g 1 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x A90030 ..

Align seg 1/1 to: A90030 from: 1 to: 14

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: gb\_pat:A90032

seq\_documentation\_block:

LOCUS A90032 14 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 213 from Patent EP0856579.  
ACCESSION A90032  
VERSION A90032.1 GI:6738546

KEYWORDS  
SOURCE  
ORGANISM  
unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 213 05-AUG-1998;  
BIODOSTIK GES (DE)

FEATURES  
source Location/Qualifiers  
1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 10 a 2 c 1 g 1 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x A90032 ..

Align seg 1/1 to: A90032 from: 1 to: 14

1 LysLysAsn 3  
|||||  
4 AAAAAAAC 12

seq\_name: gb\_pat:A90074

seq\_documentation\_block:

LOCUS A90074 14 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 255 from Patent EP0856579.  
ACCESSION A90074  
VERSION A90074.1 GI:6738588

KEYWORDS  
SOURCE  
ORGANISM  
unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 255 05-AUG-1998;  
BIODOSTIK GES (DE)

FEATURES  
source Location/Qualifiers  
1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 10 a 2 c 1 g 1 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x A90074 ..

Align seg 1/1 to: A90074 from: 1 to: 14

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: gb\_pat:A90077

seq\_documentation\_block:

LOCUS A90077 14 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 258 from Patent EP0856579.  
ACCESSION A90077  
VERSION A90077.1 GI:6738591

KEYWORDS  
SOURCE  
ORGANISM  
unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 258 05-AUG-1998;  
BIODOSTIK GES (DE)

FEATURES  
source Location/Qualifiers  
1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 9 a 2 c 1 g 2 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3



Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x A90077 ..

Align seg 1/1 to: A90077 from: 1 to: 14

1 LysLysAsn 3  
|||||  
6 AAAAAAGAAC 14

seq\_name: gb\_pat:AR031685

## seq\_documentation\_block:

LOCUS AR031685 14 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 5 from patent US 5866395.  
ACCESSION AR031685  
VERSION AR031685.1 GI:5945974  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Mathur,E.J.  
TITLE Purified thermostable pyrococcus furiosus DNA polymerase I  
JOURNAL Patent: US 5866395-A 5 02-FEB-1999;  
FEATURES  
source 1..14  
BASE COUNT 6 a 0 c 2 g 2 t 4 others  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR031685 ..

Align seg 1/1 to: AR031685 from: 1 to: 14

1 LysLysAsn 3  
|||||  
4 AAAAAAAY 12

seq\_name: gb\_pat:AR031686

## seq\_documentation\_block:

LOCUS AR031686 14 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 6 from patent US 5866395.  
ACCESSION AR031686  
VERSION AR031686.1 GI:5945975  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Mathur,E.J.  
TITLE Purified thermostable pyrococcus furiosus DNA polymerase I  
JOURNAL Patent: US 5866395-A 6 02-FEB-1999;  
FEATURES  
source 1..14  
BASE COUNT 2 a 2 c 0 g 6 t 4 others  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR031686/rev ..

Align seg 1/1 to reverse of: AR031686 from: 1 to: 14

1 LysLysAsn 3  
|||||  
11 AAAAAAAY 3

seq\_name: gb\_pat:AR051240

## seq\_documentation\_block:

LOCUS AR051240 14 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 8 from patent US 5830658.  
ACCESSION AR051240  
VERSION AR051240.1 GI:5974604  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Gryaznov,S.M.  
TITLE Convergent synthesis of branched and multiply connected  
macromolecular structures  
JOURNAL Patent: US 5830658-A 8 03-NOV-1998;  
FEATURES  
source 1..14  
BASE COUNT 12 a 2 c 0 g 0 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR051240 ..

Align seg 1/1 to: AR051240 from: 1 to: 14

1 LysLysAsn 3  
|||||  
6 AAAAAAAY 14

seq\_name: gb\_pat:AR072886

## seq\_documentation\_block:

LOCUS AR072886 14 bp DNA linear PAT 28-AUG-2000  
DEFINITION Sequence 5 from patent US 5948663.  
ACCESSION AR072886  
VERSION AR072886.1 GI:9999649  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Mathur,E.J.  
TITLE Purified thermostable pyrococcus furiosus DNA polymerase I  
JOURNAL Patent: US 5948663-A 5 07-SEP-1999;  
FEATURES  
source 1..14  
BASE COUNT 6 a 0 c 2 g 2 t 4 others  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0



Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR072886 ..

Align seg 1/1 to: AR072886 from: 1 to: 14

1 LysLysAsn 3  
|||||  
4 AAAAAAAY 12

seq\_name: gb\_pat:AR072887

## seq\_documentation\_block:

LOCUS AR072887 14 bp DNA linear PAT 28-AUG-2000  
DEFINITION Sequence 6 from patent US 5948663.  
ACCESSION AR072887  
VERSION AR072887.1 GI:9999650  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Mathur,E.J.  
TITLE Purified thermostable pyrococcus furiosus DNA polymerase I  
JOURNAL Patent: US 5948663-A 6 07-SEP-1999;  
FEATURES  
Location/Qualifiers  
1..14  
/organism="unknown"  
BASE COUNT 2 a 2 c 0 g 6 t 4 others  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR072887/rev ..

Align seg 1/1 to reverse of: AR072887 from: 1 to: 14

1 LysLysAsn 3  
|||||  
11 AAAAAAAY 3

seq\_name: gb\_pat:AR074714

## seq\_documentation\_block:

LOCUS AR074714 14 bp DNA linear PAT 28-AUG-2000  
DEFINITION Sequence 11 from patent US 5955276.  
ACCESSION AR074714  
VERSION AR074714.1 GI:10001467  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Morgante,M. and Vogel,J.Marie.  
TITLE Compound microsatellite primers for the detection of genetic polymorphisms  
JOURNAL Patent: US 5955276-A 11 21-SEP-1999;  
FEATURES  
Location/Qualifiers  
1..14  
/organism="unknown"  
BASE COUNT 0 a 2 c 2 g 10 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR074714/rev ..

Align seg 1/1 to reverse of: AR074714 from: 1 to: 14

1 LysLysAsn 3  
|||||  
12 AAAAAAAC 4

seq\_name: gb\_pat:AR119038

## seq\_documentation\_block:

LOCUS AR119038 14 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 164 from patent US 6150092.  
ACCESSION AR119038  
VERSION AR119038.1 GI:14100948  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Uchida,K., Uchida,T., Tanaka,Y., Matsuda,Y. and Kondo,S.  
TITLE Antisense nucleic acid compound targeted to VEGF  
JOURNAL Patent: US 6150092-A 164 21-NOV-2000;  
FEATURES  
Location/Qualifiers  
1..14  
/organism="unknown"  
BASE COUNT 2 a 3 c 1 g 8 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AR119038/rev ..

Align seg 1/1 to reverse of: AR119038 from: 1 to: 14

1 LysLysAsn 3  
|||||  
11 AAAAAAAT 3

seq\_name: gb\_pat:AR127787

## seq\_documentation\_block:

LOCUS AR127787 14 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 8 from patent US 6180777.  
ACCESSION AR127787  
VERSION AR127787.1 GI:14114382  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Horn,T.  
TITLE Synthesis of branched nucleic acids  
JOURNAL Patent: US 6180777-A 8 30-JAN-2001;  
FEATURES  
Location/Qualifiers  
1..14  
/organism="unknown"  
BASE COUNT 12 a 2 c 0 g 0 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000



alignment\_scores:



Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x A65238/rev ..

Align seg 1/1 to reverse of: A65238 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: gb\_pat:A88106

seq\_documentation\_block:

LOCUS A88106 15 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 254 from Patent WO9833904.  
ACCESSION A88106  
VERSION A88106.1 GI:6736676

KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 15)  
AUTHORS Brysch,W. and Schlingensiepen,K.  
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD  
JOURNAL Patent: WO 9833904-A 254 06-AUG-1998;  
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)

FEATURES  
Location/Qualifiers  
source  
1..15  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 11 a 2 c 1 g 1 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x A88106 ..

Align seg 1/1 to: A88106 from: 1 to: 15

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: gb\_pat:A88108

seq\_documentation\_block:

LOCUS A88108 15 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 256 from Patent WO9833904.  
ACCESSION A88108  
VERSION A88108.1 GI:6736678

KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 15)  
AUTHORS Brysch,W. and Schlingensiepen,K.  
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD  
JOURNAL Patent: WO 9833904-A 256 06-AUG-1998;  
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)

FEATURES  
Location/Qualifiers  
source  
1..15  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 11 a 2 c 1 g 1 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x A88108 ..

Align seg 1/1 to: A88108 from: 1 to: 15

1 LysLysAsn 3  
|||||  
5 AAAAAAAC 13

seq\_name: gb\_pat:A90073

seq\_documentation\_block:

LOCUS A90073 15 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 254 from Patent EP0856579.  
ACCESSION A90073  
VERSION A90073.1 GI:6738587

KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 15)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 254 05-AUG-1998;  
BIOGNOSTIK GES (DE)

FEATURES  
Location/Qualifiers  
source  
1..15  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 11 a 2 c 1 g 1 t  
ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x A90073 ..

Align seg 1/1 to: A90073 from: 1 to: 15

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: gb\_pat:A90075

seq\_documentation\_block:

LOCUS A90075 15 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 256 from Patent EP0856579.  
ACCESSION A90075  
VERSION A90075.1 GI:6738589

KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
unclassified.

REFERENCE 1 (bases 1 to 15)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 256 05-AUG-1998;  
BIOGNOSTIK GES (DE)

FEATURES  
Location/Qualifiers  
source  
1..15  
/organism="unidentified"



BASE COUNT	11 a	2 c	1 g	1 t	/db_xref="taxon:32644"
ORIGIN					
alignment_scores:	Quality: 16.00	Length: 3			
	Ratio: 5.333	Gaps: 0			
Percent Similarity:	100.000	Percent Identity:	100.000		
alignment_block:					
US-09-528-682-2 x A90075	..				
Align seg 1/1 to: A90075	from: 1	to: 15			
1 LysLysAsn 3					
5 AAAAAGAAC 13					
seq_name: gb_pat:AR056151					
seq_documentation_block:					
LOCUS AR056151	15 bp	DNA	linear	PAT 29-SEP-1999	
DEFINITION	Sequence 355 from patent US 5837542.				
ACCESSION	AR056151				
VERSION	AR056151.1 GI:5981728				
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 15)				
AUTHORS	Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and Draper,K.G.				
TITLE	Intercellular adhesion molecule-1 (ICAM-1) ribozymes				
JOURNAL	Patent: US 5837542-A 355 17-NOV-1998;				
FEATURES	Location/Qualifiers				
source	1..15				
BASE COUNT	2 a	0 c	1 g	12 t	
ORIGIN					
alignment_scores:	Quality: 16.00	Length: 3			
	Ratio: 5.333	Gaps: 0			
Percent Similarity:	100.000	Percent Identity:	100.000		
alignment_block:					
US-09-528-682-2 x AR056151/rev	..				
Align seg 1/1 to reverse of: AR056151	from: 1	to: 15			
1 LysLysAsn 3					
14 AAAAATAAT 6					
seq_name: gb_pat:AR056152					
seq_documentation_block:					
LOCUS AR056152	15 bp	DNA	linear	PAT 29-SEP-1999	
DEFINITION	Sequence 356 from patent US 5837542.				
ACCESSION	AR056152				
VERSION	AR056152.1 GI:5981729				
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 15)				
AUTHORS	Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and Draper,K.G.				
TITLE	Intercellular adhesion molecule-1 (ICAM-1) ribozymes				
JOURNAL	Patent: US 5837542-A 356 17-NOV-1998;				
FEATURES	Location/Qualifiers				
source	1..15				
BASE COUNT	2 a	0 c	1 g	12 t	
ORIGIN					
alignment_scores:	Quality: 16.00	Length: 3			
	Ratio: 5.333	Gaps: 0			
Percent Similarity:	100.000	Percent Identity:	100.000		
alignment_block:					
US-09-528-682-2 x AR056151/rev	..				
Align seg 1/1 to reverse of: AR056151	from: 1	to: 15			
1 LysLysAsn 3					
14 AAAAATAAT 6					
seq_name: gb_pat:AR056152					
seq_documentation_block:					
LOCUS AR056152	15 bp	DNA	linear	PAT 29-SEP-1999	
DEFINITION	Sequence 356 from patent US 5837542.				
ACCESSION	AR056152				
VERSION	AR056152.1 GI:5981729				
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 15)				
AUTHORS	Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and Draper,K.G.				
TITLE	Intercellular adhesion molecule-1 (ICAM-1) ribozymes				
JOURNAL	Patent: US 5837542-A 356 17-NOV-1998;				
FEATURES	Location/Qualifiers				
source	1..15				
BASE COUNT	2 a	0 c	1 g	12 t	
ORIGIN					
alignment_scores:	Quality: 16.00	Length: 3			
	Ratio: 5.333	Gaps: 0			
Percent Similarity:	100.000	Percent Identity:	100.000		
alignment_block:					
US-09-528-682-2 x AR056151/rev	..				
Align seg 1/1 to reverse of: AR056151	from: 1	to: 15			
1 LysLysAsn 3					
14 AAAAATAAT 6					
seq_name: gb_pat:AR056152					
seq_documentation_block:					
LOCUS AR056152	15 bp	DNA	linear	PAT 29-SEP-1999	
DEFINITION	Sequence 356 from patent US 5837542.				
ACCESSION	AR056152				
VERSION	AR056152.1 GI:5981729				
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 15)				
AUTHORS	Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and Draper,K.G.				
TITLE	Intercellular adhesion molecule-1 (ICAM-1) ribozymes				
JOURNAL	Patent: US 5837542-A 356 17-NOV-1998;				
FEATURES	Location/Qualifiers				
source	1..15				
BASE COUNT	2 a	0 c	1 g	12 t	
ORIGIN					
alignment_scores:	Quality: 16.00	Length: 3			
	Ratio: 5.333	Gaps: 0			



FEATURES                      Location/Qualifiers  
1..15  
source                      1 a                      0 c                      1 g                      13 t  
BASE COUNT  
ORIGIN

alignment\_scores:  
Quality: 16.00                      Length: 3  
Ratio: 5.333                      Gaps: 0  
Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AR056154/rev ..

Align seg 1/1 to reverse of: AR056154 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAT 3

seq\_name: gb\_pat:AR056155

seq\_documentation\_block:

LOCUS AR056155                      15 bp                      DNA                      linear                      PAT 29-SEP-1999

DEFINITION Sequence 359 from patent US 5837542.

ACCESSION AR056155

VERSION AR056155.1 GI:5981732

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 15)

AUTHORS

Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and

Draper,K.G.

Interleukin adhesion molecule-1 (ICAM-1) ribozymes

JOURNAL

Patent: US 5837542-A 359 17-NOV-1998;

FEATURES

Location/Qualifiers

1..15

source

BASE COUNT

1 a                      0 c                      1 g                      13 t

ORIGIN

alignment\_scores:

Quality: 16.00                      Length: 3

Ratio: 5.333                      Gaps: 0

Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR056155/rev ..

Align seg 1/1 to reverse of: AR056155 from: 1 to: 15

1 LysLysAsn 3  
|||||  
10 AAAAAAAT 2

seq\_name: gb\_pat:AR056156

seq\_documentation\_block:

LOCUS AR056156                      15 bp                      DNA                      linear                      PAT 29-SEP-1999

DEFINITION Sequence 360 from patent US 5837542.

ACCESSION AR056156

VERSION AR056156.1 GI:5981733

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 15)

AUTHORS

Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and

Draper,K.G.

Interleukin adhesion molecule-1 (ICAM-1) ribozymes

JOURNAL

Patent: US 5837542-A 360 17-NOV-1998;

FEATURES

Location/Qualifiers

1..15

source

JOURNAL Patent: US 5837542-A 360 17-NOV-1998;  
FEATURES                      Location/Qualifiers  
1..15  
source                      1 a                      0 c                      0 g                      14 t  
BASE COUNT  
ORIGIN

alignment\_scores:  
Quality: 16.00                      Length: 3  
Ratio: 5.333                      Gaps: 0  
Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR056156/rev ..

Align seg 1/1 to reverse of: AR056156 from: 1 to: 15

1 LysLysAsn 3  
|||||  
9 AAAAAAAT 1

seq\_name: gb\_pat:AR079608

seq\_documentation\_block:

LOCUS AR079608                      15 bp                      DNA                      linear                      PAT 31-AUG-2000

DEFINITION Sequence 3 from patent US 5965721.

ACCESSION AR079608

VERSION AR079608.1 GI:10006352

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 15)

AUTHORS

Cook,P.Dan, Sanghvi,Y.Shantilal, Vasseur,J.Jacques and Debart,F.

Interleukin adhesion molecule-1 (ICAM-1) ribozymes

JOURNAL

Patent: US 5965721-A 3 12-OCT-1999;

FEATURES

Location/Qualifiers

1..15

source

BASE COUNT

0 a                      2 c                      3 g                      10 t

ORIGIN

alignment\_scores:

Quality: 16.00                      Length: 3

Ratio: 5.333                      Gaps: 0

Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR079608/rev ..

Align seg 1/1 to reverse of: AR079608 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: gb\_pat:AR113909

seq\_documentation\_block:

LOCUS AR113909                      15 bp                      DNA                      linear                      PAT 16-MAY-2001

DEFINITION Sequence 355 from patent US 6132967.

ACCESSION AR113909

VERSION AR113909.1 GI:14094231

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 15)

AUTHORS

Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and

Draper,K.G.

Interleukin adhesion molecule-1 (ICAM-1) ribozymes

JOURNAL

Patent: US 6132967-A 355 16-MAY-2001;

FEATURES

Location/Qualifiers

1..15

source

BASE COUNT

0 a                      2 c                      3 g                      10 t

ORIGIN

alignment\_scores:

Quality: 16.00                      Length: 3

Ratio: 5.333                      Gaps: 0

Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AR079608/rev ..

Align seg 1/1 to reverse of: AR079608 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: gb\_pat:AR113909

seq\_documentation\_block:

LOCUS AR113909                      15 bp                      DNA                      linear                      PAT 16-MAY-2001

DEFINITION Sequence 355 from patent US 6132967.

ACCESSION AR113909

VERSION AR113909.1 GI:14094231

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 15)

AUTHORS

Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and

Draper,K.G.

Interleukin adhesion molecule-1 (ICAM-1) ribozymes

JOURNAL

Patent: US 6132967-A 355 16-MAY-2001;

FEATURES

Location/Qualifiers

1..15

source

BASE COUNT

0 a                      2 c                      3 g                      10 t

ORIGIN



intercellular adhesion molecule-1 (ICAM-1)  
Patent: US 6132967-A 355 17-OCT-2000;  
JOURNAL  
FEATURES  
source  
1. .15  
BASE COUNT 2 a 0 c 1 g 12 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AR113909/rev ..

Align seg 1/1 to reverse of: AR113909 from: 1 to: 15

1 LysLysAsn 3  
|||||  
14 AAAAAAAT 6

seq\_name: gb\_pat:AR113910

seq\_documentation\_block:  
LOCUS AR113910 15 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 356 from patent US 6132967.

ACCESSION AR113910  
VERSION AR113910.1 GI:14094232

KEYWORDS  
SOURCE  
ORGANISM  
Unknown.

REFERENCE 1 (bases 1 to 15)

AUTHORS  
Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and  
Draper,K.G.

TITLE  
Ribozyme treatment of diseases or conditions related to levels of  
intercellular adhesion molecule-1 (ICAM-1)

JOURNAL  
FEATURES  
source  
1. .15  
BASE COUNT 1 a 0 c 1 g 13 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AR113910/rev ..

Align seg 1/1 to reverse of: AR113910 from: 1 to: 15

1 LysLysAsn 3  
|||||  
13 AAAAAAAT 5

seq\_name: gb\_pat:AR113911

seq\_documentation\_block:  
LOCUS AR113911 15 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 357 from patent US 6132967.

ACCESSION AR113911  
VERSION AR113911.1 GI:14094233

KEYWORDS  
SOURCE  
ORGANISM  
Unknown.

REFERENCE 1 (bases 1 to 15)

AUTHORS  
Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and  
Draper,K.G.

TITLE  
Ribozyme treatment of diseases or conditions related to levels of  
intercellular adhesion molecule-1 (ICAM-1)

JOURNAL  
FEATURES  
source  
1. .15  
BASE COUNT 1 a 0 c 1 g 13 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AR113911/rev ..

Align seg 1/1 to reverse of: AR113911 from: 1 to: 15

1 LysLysAsn 3  
|||||  
12 AAAAAAAT 4

seq\_name: gb\_pat:AR113912

seq\_documentation\_block:  
LOCUS AR113912 15 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 358 from patent US 6132967.

ACCESSION AR113912  
VERSION AR113912.1 GI:14094234

KEYWORDS  
SOURCE  
ORGANISM  
Unknown.

REFERENCE 1 (bases 1 to 15)

AUTHORS  
Grimm,S., Stinchcomb,D.T., McSwiggen,J., Sullivan,S. and  
Draper,K.G.

TITLE  
Ribozyme treatment of diseases or conditions related to levels of  
intercellular adhesion molecule-1 (ICAM-1)

JOURNAL  
FEATURES  
source  
1. .15  
BASE COUNT 1 a 0 c 1 g 13 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AR113912/rev ..

Align seg 1/1 to reverse of: AR113912 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAT 3











OM of: US-09-528-682-2 to: N\_Geneseq\_032802:\* out\_format : pfs

Date: Jun 18, 2002 7:39 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=framet.p2n.model -DEV=xlp  
-Q/cg2\_1/USPTO.spool/US09528682/runat\_18062002\_082444\_7843/app\_query.fasta\_1.689  
-DB=N\_Geneseq\_032802 -QPM=fastap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -GAPEXT=0.500  
-GAPOP=6.000 -GAPEXT=7.000 -GAPOP=10.000 -GAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bclosum62  
-TRANS=human40.cdi -LISP=1000 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFM=pfs  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USDS=US09528682.ecgn1\_1.677 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-2

Query length: 3

Database: N\_Geneseq\_032802:\*

Database sequences: 1736436

Database length: 85845721

Search time (sec): 554.690000

score\_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV50073	10	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA86504	10	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF29870	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF33290	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF34095	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF34588	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF35488	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF36874	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF37316	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF38397	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF38535	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF40328	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF40495	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF41776	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF42242	1	102.39	2.3e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH19036	1	101.70	2.6e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV73560	12	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA943316	12	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA103356	12	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH46026	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67135	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67157	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67218	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67278	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67356	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67386	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67428	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67458	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67482	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67505	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67640	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67643	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67751	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67810	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67912	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67913	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67928	1	101.07	2.8e+03	16.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68005	1	101.07	2.8e+03	16.00	

/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68036	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68090	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68175	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68195	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68219	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68243	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68246	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68268	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68290	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68346	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68430	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68452	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68459	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68490	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68505	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68574	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68575	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68665	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68690	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68732	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68733	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68751	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68792	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68911	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68963	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69045	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69099	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69177	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69383	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69403	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69518	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69571	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69623	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69666	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69755	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69809	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69855	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69886	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH69898	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70043	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70135	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70136	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70137	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70325	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70329	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70609	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70757	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70820	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70843	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70868	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70893	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH70899	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71004	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71052	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71240	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71292	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71297	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71359	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71401	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71441	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71442	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71473	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71622	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH71699	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72193	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72249	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72262	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72306	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72500	+	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72501	-	16.00	101.07	2.8e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH72568	-	16.00	101.07	2.8e+03



























/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC56537 + 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC56950 - 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC56951 + 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57548 - 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57549 + 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57630 - 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57631 + 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57956 - 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC57957 + 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC58202 - 16.00 100.49 3.0e+03  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABC58203 + 16.00 100.49 3.0e+03

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV50073

seq\_documentation\_block:

ID AAV50073 standard; DNA; 10 BP.

XX AAV50073;

DT 21-OCT-1998 (first entry)

XX Yeast tag for highly expressed gene PYK1.

XX Yeast; Saccharomyces cerevisiae; transcriptome; cell cycle;

KW regulation; eukaryotic cell; antifungal; SAGE tag; gene expression;

KW serial analysis of gene expression; probe; ss.

XX Saccharomyces cerevisiae.

OS Synthetic.

XX WO9832847-A2.

XX 30-JUL-1998.

XX 22-JAN-1998; 98WO-US01216.

XX 23-JAN-1997; 97US-0035917.

XX (UWJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX Kinzler KW, Velculescu VE, Vogelstein B;

XX WPI; 1998-427943/36.

XX Yeast transcriptome - useful for modulating eukaryotic cell, for  
PT screening antifungal agents, and for identifying genes in cell cycle  
PT progression

XX Claim 11; Page 21; 44pp; English.

XX Yeast transcriptome is encoded by a DNA molecule comprising a yeast  
CC gene involved in cell cycle progression selected from the group of  
CC nonannotated ORF (NORF) genes. SAGE (serial analysis gene expression)  
CC tags for highly expressed genes and NORF genes are given in AAV50051 to  
CC AAV50345. The present invention describes: (1) a method of using yeast  
CC genes to modulate the cell cycle which comprises administering to a cell  
CC an isolated DNA molecule comprising a yeast gene which is involved in  
CC cell cycle progression selected from differentially expressed genes  
CC (SAGE tags given in AAV50051 to AAV50345); (2) a method for screening  
CC candidate antifungal drugs which comprises contacting a test substance  
CC with a yeast cell and monitoring expression of a yeast gene which is  
CC involved in cell cycle progression; (3) a method of identifying human  
CC genes which are involved in cell cycle progression which comprises  
CC hybridizing a probe comprising at least 10 contiguous nucleotides of a  
CC yeast gene which is differentially expressed between at least 2 phases  
CC selected from the log phase, the S phase and the G2/M phase; and (4) a  
CC probe for ascertaining the phase in the cell cycle, where the probe  
CC comprises at least 14 contiguous nucleotides of a NORF gene (SAGE tags  
CC given in AAV50051 to AAV50345), or as an array of probes on a solid  
CC support.

XX Sequence 10 BP; 0 A; 1 C; 1 G; 8 T; 0 other;

50

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AAV50073/rev ..

Align seg 1/1 to reverse of: AAV50073 from: 1 to: 10

1 LysLysAsn 3

|||||

9 AAGAAAAAC 1

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ86504

seq\_documentation\_block:

ID AAZ86504 standard; DNA; 10 BP.

XX AAZ86504;

XX 07-APR-2000 (first entry)

XX Metastatic breast tumour cell downregulated transcript tag #5738.

XX Human; metastatic breast tumour tissue; breast cancer; tag; primer;

KW non-metastatic breast tumour tissue; gene therapy; anticancer;

KW antimetastatic; vaccine; diagnosis; ss.

XX Homo sapiens.

XX WO9965928-A2.

XX 23-DEC-1999.

XX 18-JUN-1999; 99WO-US13647.

XX 19-JUN-1998; 98US-0089853.

PR 19-JUN-1998; 98US-0089997.

PR 19-JUN-1998; 98US-0090039.

PR 19-JUN-1998; 98US-0090040.

XX 19-JUN-1998; 98US-0090041.

(GENZ ) GENZYME CORP.

(ROBE/) ROBERTS B L.

(SHAN/) SHANKARA S.

XX Roberts BL, Shankara S;

XX WPI; 2000-106079/09.

XX Isolated polynucleotides differentially expressed between metastatic  
PT and non-metastatic breast cancer cells, useful for diagnosis,  
PT prevention and treatment of cancer -

XX Claim 1; Page 210; 219pp; English.

XX AAZ80767 to AAZ83941 represent tags corresponding to distinct  
CC transcripts that are preferentially transcribed in the metastatic breast  
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
CC AAZ83942 to AAZ86677 represent tags corresponding to distinct transcripts  
CC that are preferentially transcribed in the primary or non-metastatic  
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
CC cells). These transcripts can be used for diagnosis, prognosis, metastatic  
CC monitoring and treatment of breast cancer, particularly where metastatic.  
CC diagnosis is by standard immunoassays or hybridisation/amplification  
CC reactions. Compounds that modulate expression of the transcripts are  
CC potentially useful for treatment of (metastatic) breast cancer, while  
CC promoters from the transcripts are used to direct expression, in selected  
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
CC sequences), particularly an antigen-encoding sequence for use in gene or  
CC cell-based vaccines. Polypeptides encoded by the transcripts are also



CC useful in vaccines; for diagnosing breast cancer and for raising  
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
CC therapeutic agents. Host cells that produce the polypeptides can be used  
CC to expand and isolate populations of educated, antigen-specific immune  
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
CC adoptive immunotherapy.

XX Sequence 10 BP; 2 A; 1 C; 0 G; 7 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AA286504/rev ..

Align seg 1/1 to reverse of: AA286504 from: 1 to: 10

1 LysLysAsn 3  
|||||  
9 AAAAAAAT 1

seq\_name: /SID85/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF29870

seq\_documentation\_block:  
ID AAF29870 standard; DNA; 10 BP.

XX AC AAF29870;  
XX DT 03-APR-2001 (first entry)  
XX DE Starting 1 fragment #3.

XX KW Cloning; exon shuffling; store; adapter; ss.

XX OS Unidentified.

XX PN WO200100816-A1.

XX PD 04-JAN-2001.

XX PF 27-JUN-2000; 2000WO-GB02512.

XX PR 28-JUN-1999; 99NO-0001325.

XX PR 20-JUN-2000; 2000NO-0003190.

XX PR 20-JUN-2000; 2000NO-0003191.

XX PA (COMP-) COMPLETE GENOMICS AS.

XX PI (JONE/) JONES E L.

XX PI Lexow P;

XX DR WPI; 2001-123006/13.

XX PT Attaching fragments of first nucleic acids to second nucleic acids by  
XX use of adapters complementary to first single stranded regions on the  
XX first molecules but which have a different single stranded region at  
XX the other terminus -

XX PS Disclosure; Fig 2; 100pp; English.

XX CC The present invention relates to attaching a fragment of first and  
XX second nucleic acid molecules involves use of an adapter molecule  
XX which is complementary to a single stranded region generated on the  
XX target but which has a different single stranded region at its  
XX other terminus and therefore modifies single stranded regions  
XX presented for binding by the target. Attaching first and second  
XX nucleic acid molecules may be used in cloning. The method can  
XX also be used for exon shuffling other recombinations that are  
XX relevant in connection with artificial evolutionary systems.  
XX The fragment chains may be used to store information.

XX SQ Sequence 10 BP; 9 A; 1 C; 0 G; 0 U; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAF29870 ..

Align seg 1/1 to: AAF29870 from: 1 to: 10

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: /SID85/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF33290

seq\_documentation\_block:  
ID AAF33290 standard; DNA; 10 BP.

XX AC AAF33290;

XX DT 23-MAR-2001 (first entry)

XX DE Yeast highly expressed gene SAGE tag oligonucleotide SEQ ID NO:29.

XX KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
XX nor previously assigned open reading frame; nonannotated ORF; SAGE;  
XX serial analysis of gene expression; antifungal; tag; identification;  
XX linker; PCR primer; ds.

XX OS Saccharomyces cerevisiae.

XX PN WO200077214-A2.

XX PD 21-DEC-2000.

XX PF 14-JUN-2000; 2000WO-US16223.

XX PR 16-JUN-1999; 99US-0335032.

XX PA (UYJO ) UNIV JOHNS HOPKINS.

XX PI Velculescu V, Vogelstein B, Kinzler K;

XX DR WPI; 2001-061874/07.

XX PT Yeast gene coding sequences comprising NORF genes with serial analysis  
XX of gene expression (SAGE) tags, useful for studying, monitoring and  
XX affecting phases of the cell cycle -

XX PS Example; Page 21; 419pp; English.

XX CC The present invention describes an isolated DNA molecule comprising a  
XX coding sequence of a yeast gene selected from a group of 745 NORF (not  
XX previously assigned open reading frame; or nonannotated ORF) genes  
XX comprising a SAGE (serial analysis of gene expression) tag. Also  
XX described are: (1) a method (M1) of using NORF genes to affect the cell  
XX cycle comprising administering a NORF gene whose expression varies by at  
XX least 10% between any two phases of the cell cycle selected from log  
XX phase, S phase and G2/M; (2) a method (M2) for screening candidate  
XX antifungal drugs comprising: (a) contacting a test substance with a  
XX yeast cell; and (b) monitoring expression of a NORF gene whose  
XX expression varies as in M1, where a test substance which modifies the  
XX expression of the yeast gene is a candidate antifungal drug; (3) a method  
XX (M3) for identifying human genes which are involved in cell cycle  
XX progression comprising contacting human DNA with a probe which comprises  
XX at least 10 contiguous nucleotides of a NORF gene whose expression varies  
XX as in M1; and (4) a method (M4) for identifying a candidate drug as a  
XX member of a class of drugs having a characteristic effect on gene



CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 10 BP; 0 A; 1 C; 1 G; 8 T; 0 other;

alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-528-682-2 x AAF33290/rev ..

Align seg 1/1 to reverse of: AAF33290 from: 1 to: 10

1 LysLysAsn 3  
 |||||  
 9 AAGAAAAAC 1

seq\_name: /STD5/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF34095

seq\_documentation\_block:  
 ID AAF34095 standard; DNA; 10 BP.

AC AAF34095;

DT 23-MAR-2001 (first entry)

DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:834.

XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
 KW serial analysis of gene expression; antifungal; tag; identification;  
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO200077214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI; 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis  
 PT of gene expression (SAGE) tags, useful for studying, monitoring and  
 PT affecting phases of the cell cycle -

XX Example; Page 29; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a  
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
 CC previously assigned open reading frame; or nonannotated ORF) genes  
 CC comprising a SAGE (serial analysis of gene expression) tag. Also  
 CC described are: (1) a method (M1) of using NORF genes to affect the cell  
 CC cycle comprising administering a NORF gene whose expression varies by at

CC least 10% between any two phases of the cell cycle selected from log  
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
 CC antifungal drugs comprising: (a) contacting a test substance with a  
 CC yeast cell; and (b) monitoring expression of a NORF gene whose  
 CC expression varies as in M1, where a test substance which modifies the  
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
 CC (M3) for identifying human genes which are involved in cell cycle  
 CC progression comprising contacting human DNA with a probe which comprises  
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
 CC member of a class of drugs having a characteristic effect on gene  
 CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 10 BP; 0 A; 1 C; 1 G; 8 T; 0 other;

alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-528-682-2 x AAF34095/rev ..

Align seg 1/1 to reverse of: AAF34095 from: 1 to: 10

1 LysLysAsn 3  
 |||||  
 9 AAGAAAAAC 1

seq\_name: /STD5/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF34588

seq\_documentation\_block:  
 ID AAF34588 standard; DNA; 10 BP.

XX AAF34588;

XX 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:1327.

XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
 KW serial analysis of gene expression; antifungal; tag; identification;  
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO200077214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI; 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis



PT of gene expression (SAGE) tags, useful for studying, monitoring and  
 PT affecting phases of the cell cycle -  
 PS Example; Page 47; 419pp; English.  
 XX  
 CC The present invention describes an isolated DNA molecule comprising a  
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
 CC previously assigned open reading frame; or nonannotated ORF) genes  
 CC comprising a SAGE (serial analysis of gene expression) tag. Also  
 CC described are: (1) a method (M1) of using NORF genes to affect the cell  
 CC cycle comprising administering a NORF gene whose expression varies by at  
 CC least 10% between any two phases of the cell cycle selected from log  
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
 CC antifungal drugs comprising: (a) contacting a test substance with a  
 CC yeast cell; and (b) monitoring expression of a NORF gene whose  
 CC expression varies as in M1, where a test substance which modifies the  
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
 CC (M3) for identifying human genes which are involved in cell cycle  
 CC progression comprising contacting human DNA with a probe which comprises  
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
 CC member of a class of drugs having a characteristic effect on gene  
 CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 10 BP; 1 A; 3 C; 0 G; 6 T; 0 other;

alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-528-682-2 x AAF34588/rev ..

Align seg 1/1 to reverse of: AAF34588 from: 1 to: 10

1 LysLysAsn 3  
 |||||  
 10 AAGAAGAAT 2

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF35488

seq\_documentation\_block:  
 ID AAF35488 standard; DNA; 10 BP.  
 XX  
 AC AAF35488;  
 XX  
 DT 23-MAR-2001 (first entry)  
 XX  
 DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:2227.  
 XX  
 KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
 KW serial analysis of gene expression; antifungal; tag; identification;  
 KW linker; PCR primer; ds.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200077214-A2.  
 XX  
 PD 21-DEC-2000.  
 XX

PF 14-JUN-2000; 2000WO-US16223.  
 XX  
 PR 16-JUN-1999; 99US-0335032.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Velculescu V, Vogelstein B, Kinzler K;  
 XX  
 DR WPI; 2001-061874/07.  
 XX  
 PT Yeast gene coding sequences comprising NORF genes with serial analysis  
 PT of gene expression (SAGE) tags, useful for studying, monitoring and  
 PT affecting phases of the cell cycle -  
 PS Example; Page 79; 419pp; English.  
 XX  
 CC The present invention describes an isolated DNA molecule comprising a  
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
 CC previously assigned open reading frame; or nonannotated ORF) genes  
 CC comprising a SAGE (serial analysis of gene expression) tag. Also  
 CC described are: (1) a method (M1) of using NORF genes to affect the cell  
 CC cycle comprising administering a NORF gene whose expression varies by at  
 CC least 10% between any two phases of the cell cycle selected from log  
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
 CC antifungal drugs comprising: (a) contacting a test substance with a  
 CC yeast cell; and (b) monitoring expression of a NORF gene whose  
 CC expression varies as in M1, where a test substance which modifies the  
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
 CC (M3) for identifying human genes which are involved in cell cycle  
 CC progression comprising contacting human DNA with a probe which comprises  
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
 CC member of a class of drugs having a characteristic effect on gene  
 CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 10 BP; 1 A; 2 C; 1 G; 6 T; 0 other;

alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-528-682-2 x AAF35488/rev ..

Align seg 1/1 to reverse of: AAF35488 from: 1 to: 10

1 LysLysAsn 3  
 |||||  
 10 AAGAAGAAT 2

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF36874

seq\_documentation\_block:  
 ID AAF36874 standard; DNA; 10 BP.  
 XX  
 AC AAF36874;  
 XX  
 DT 23-MAR-2001 (first entry)  
 XX  
 DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:3613.  
 XX



KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.  
XX  
XX Saccharomyces cerevisiae.  
XX WO200077214-A2.  
XX 21-DEC-2000.  
XX  
XX 14-JUN-2000; 2000WO-US16223.  
XX  
XX 16-JUN-1999; 99US-0335032.  
XX (UYJO ) UNIV JOHNS HOPKINS.  
XX  
XX Velculescu V, Vogelstein B, Kinzler K;  
XX WPI; 2001-061874/07.  
XX  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 129; 419pp; English.  
XX  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 10 BP; 9 A; 1 C; 0 G; 0 U; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAF36874 ..

Align seg 1/1 to: AAF36874 from: 1 to: 10

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: /STDS5/gcdata/geneseq/geneseq-embl/NA2001A.DAT:AAF37316  
seq\_documentation\_block:  
ID AAF37316 standard; DNA; 10 BP.  
XX  
XX AAF37316;  
XX  
XX 23-MAR-2001 (first entry)  
XX  
XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:4055.  
XX  
XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.  
XX  
XX Saccharomyces cerevisiae.  
XX WO200077214-A2.  
XX 21-DEC-2000.  
XX  
XX 14-JUN-2000; 2000WO-US16223.  
XX  
XX 16-JUN-1999; 99US-0335032.  
XX (UYJO ) UNIV JOHNS HOPKINS.  
XX  
XX Velculescu V, Vogelstein B, Kinzler K;  
XX WPI; 2001-061874/07.  
XX  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 144; 419pp; English.  
XX  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 10 BP; 1 A; 0 C; 0 G; 9 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0



Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AAF37316/rev ..

Align seg 1/1 to reverse of: AAF37316 from: 1 to: 10

```
1 LysLysAsn 3
|||||
9 AAAAAAAT 1
```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF38396

# seq\_documentation\_block:

ID AAF38396 standard; DNA; 10 BP.

XX AC AAF38396;

XX DT 23-MAR-2001 (first entry)

XX DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:5135.

XX KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

XX OS Saccharomyces cerevisiae.

XX PN WO200077214-A2.

XX PD 21-DEC-2000.

XX PF 14-JUN-2000; 2000WO-US16223.

XX PR 16-JUN-1999; 99US-0335032.

XX PA (UVJO ) UNIV JOHNS HOPKINS.

XX PI Velculescu V, Vogelstein B, Kinzler K;

XX DR WPT; 2001-061874/07.

XX PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

XX PS Example; Page 183; 419pp; English.

XX CC The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.

CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.

XX SQ Sequence 10 BP; 9 A; 0 C; 0 G; 1 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AAF38396 ..

Align seg 1/1 to: AAF38396 from: 1 to: 10

```
1 LysLysAsn 3
|||||
2 AAAAAAAT 10
```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF38397

# seq\_documentation\_block:

ID AAF38397 standard; DNA; 10 BP.

XX AC AAF38397;

XX DT 23-MAR-2001 (first entry)

XX DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:5136.

XX KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

XX OS Saccharomyces cerevisiae.

XX PN WO200077214-A2.

XX PD 21-DEC-2000.

XX PF 14-JUN-2000; 2000WO-US16223.

XX PR 16-JUN-1999; 99US-0335032.

XX PA (UVJO ) UNIV JOHNS HOPKINS.

XX PI Velculescu V, Vogelstein B, Kinzler K;

XX DR WPT; 2001-061874/07.

XX PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

XX PS Example; Page 183; 419pp; English.

XX CC The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle







```

XX DR WPI; 2001-061874/07.
XX
XX
XX PT Yeast gene coding sequences comprising NORF genes with serial analysis
XX of gene expression (SAGE) tags, useful for studying, monitoring and
XX affecting phases of the cell cycle.
XX
XX PS Example; Page 252; 419pp; English.
XX
XX CC The present invention describes an isolated DNA molecule comprising a
XX coding sequence of a yeast gene selected from a group of 745 NORF (not
XX previously assigned open reading frame; or nonannotated ORF) genes
XX comprising a SAGE (serial analysis of gene expression) tag. Also
XX described are: (1) a method (M1) of using NORF genes to affect the cell
XX cycle comprising administering a NORF gene whose expression varies by at
XX least 10% between any two phases of the cell cycle selected from log
XX phase, S phase and G2/M; (2) a method (M2) for screening candidate
XX antifungal drugs comprising: (a) contacting a test substance with a
XX yeast cell; and (b) monitoring expression of a NORF gene whose
XX expression varies as in M1, where a test substance which modifies the
XX expression of the yeast gene is a candidate antifungal drug; (3) a method
XX (M3) for identifying human genes which are involved in cell cycle
XX progression comprising contacting human DNA with a probe which comprises
XX at least 10 contiguous nucleotides of a NORF gene whose expression varies
XX as in M1; and (4) a method (M4) for identifying a candidate drug as a
XX member of a class of drugs having a characteristic effect on gene
XX expression in a yeast cell comprising contacting a yeast cell with a
XX candidate drug and monitoring expression in the yeast cell of at least 1
XX NORF gene whose expression is affected by the class of drugs. The NORF
XX genes may be used to study, monitor and affect phases of the cell cycle,
XX the differentially expressed genes may be used as markers of phases of
XX the cell cycle. The methods may be used to identify candidate drugs which
XX affect the cell cycle and for identification of antifungal drugs.
XX CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
XX the present invention. AAF33282 to AAF33267 represent linkers and PCR
XX primers used in the SAGE method, in the exemplification of the present
XX invention.
XX
XX SQ Sequence 10 BP; 1 A; 1 C; 0 G; 8 T; 0 other;

alignment_scores:
    Quality: 16.00 Length: 3
    Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AAF40328/rev ..

Align seg 1/1 to reverse of: AAF40328 from: 1 to: 10

    1 LysLysAsn 3
    |||||
    9 AAAAAAAT 1

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF40495
seq_documentation_block:
ID AAF40495 standard; DNA; 10 BP.
XX
XX AC AAF40495;
XX
XX DT 23-MAR-2001 (first entry)
XX
XX DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO: 7234.
XX
XX KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
XX nor previously assigned open reading frame; nonannotated ORF; SAGE;
XX serial analysis of gene expression; antifungal; tag; identification;
XX linker; PCR primer; ds.
XX
XX OS Saccharomyces cerevisiae.
XX

```

```

PN WO200077214-A2.
XX
XX PD 21-DEC-2000.
XX
XX PF 14-JUN-2000; 2000WO-US16223.
XX
XX PR 16-JUN-1999; 99US-0335032.
XX
XX PA (UWJO ) UNIV JOHNS HOPKINS.
XX
XX PI Veiculescu V, Vogelstein B, Kinzler K;
XX DR WPI; 2001-061874/07.
XX
XX PT Yeast gene coding sequences comprising NORF genes with serial analysis
XX of gene expression (SAGE) tags, useful for studying, monitoring and
XX affecting phases of the cell cycle.
XX
XX PS Example; Page 258; 419pp; English.
XX
XX CC The present invention describes an isolated DNA molecule comprising a
XX coding sequence of a yeast gene selected from a group of 745 NORF (not
XX previously assigned open reading frame; or nonannotated ORF) genes
XX comprising a SAGE (serial analysis of gene expression) tag. Also
XX described are: (1) a method (M1) of using NORF genes to affect the cell
XX cycle comprising administering a NORF gene whose expression varies by at
XX least 10% between any two phases of the cell cycle selected from log
XX phase, S phase and G2/M; (2) a method (M2) for screening candidate
XX antifungal drugs comprising: (a) contacting a test substance with a
XX yeast cell; and (b) monitoring expression of a NORF gene whose
XX expression varies as in M1, where a test substance which modifies the
XX expression of the yeast gene is a candidate antifungal drug; (3) a method
XX (M3) for identifying human genes which are involved in cell cycle
XX progression comprising contacting human DNA with a probe which comprises
XX at least 10 contiguous nucleotides of a NORF gene whose expression varies
XX as in M1; and (4) a method (M4) for identifying a candidate drug as a
XX member of a class of drugs having a characteristic effect on gene
XX expression in a yeast cell comprising contacting a yeast cell with a
XX candidate drug and monitoring expression in the yeast cell of at least 1
XX NORF gene whose expression is affected by the class of drugs. The NORF
XX genes may be used to study, monitor and affect phases of the cell cycle,
XX the differentially expressed genes may be used as markers of phases of
XX the cell cycle. The methods may be used to identify candidate drugs which
XX affect the cell cycle and for identification of antifungal drugs.
XX CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
XX the present invention. AAF33282 to AAF33267 represent linkers and PCR
XX primers used in the SAGE method, in the exemplification of the present
XX invention.
XX
XX SQ Sequence 10 BP; 7 A; 1 C; 1 G; 1 T; 0 other;

alignment_scores:
    Quality: 16.00 Length: 3
    Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AAF40495 ..

Align seg 1/1 to: AAF40495 from: 1 to: 10

    1 LysLysAsn 3
    |||||
    2 AAAAAAAT 10

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF41776
seq_documentation_block:
ID AAF41776 standard; DNA; 10 BP.
XX
XX AC AAF41776;
XX

```



DT 23-MAR-2001 (first entry)  
XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:8515.  
XX  
KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.  
XX  
XX Saccharomyces cerevisiae.  
XX  
XX W0200077214-A2.  
XX  
XX 21-DEC-2000.  
XX  
XX 14-JUN-2000; 2000WO-US16223.  
XX  
XX 16-JUN-1999; 99US-0335032.  
XX  
XX (UYJO ) UNIV JOHNS HOPKINS.  
XX  
XX Velculescu V, Vogelstein B, Kinzler K;  
XX WPI; 2001-061874/07.  
XX  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 304; 419pp; English.  
XX  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF4064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
XX invention.  
XX Sequence 10 BP; 6 A; 1 C; 3 G; 0 U; 0 other;  
SQ

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAF41776 ..  
Align seg 1/1 to: AAF41776 from: 1 to: 10

1 LysLysAsn 3  
|||||||  
2 AAGAAGAAC 10  
seq\_name: /STD5/gcdata/geneseq/geneseq-embl/NA2001A.DAT:AAF42242  
seq\_documentation\_block:  
ID AAF42242 standard; DNA; 10 BP.  
XX  
XX AC AAF42242;  
XX  
XX DT 23-MAR-2001 (first entry)  
XX  
XX DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:8981.  
XX  
XX KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.  
XX  
XX OS Saccharomyces cerevisiae.  
XX  
XX PN W0200077214-A2.  
XX  
XX PD 21-DEC-2000.  
XX  
XX PF 14-JUN-2000; 2000WO-US16223.  
XX  
XX PR 16-JUN-1999; 99US-0335032.  
XX  
XX PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
XX PI Velculescu V, Vogelstein B, Kinzler K;  
XX WPI; 2001-061874/07.  
XX  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 320; 419pp; English.  
XX  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF4064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
XX invention.  
XX Sequence 10 BP; 1 A; 1 C; 1 G; 7 T; 0 other;  
SQ



alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAF42242/rev ..

Align seg 1/1 to reverse of: AAF42242 from: 1 to: 10

1 LysLysAsn 3  
|||||  
10 AAGAAAAAT 2

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH19036

seq\_documentation\_block:  
ID AAH19036 standard; DNA; 11 BP.

XX  
AC AAH19036;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Diagnostic variant primer #3.  
XX  
KW CTER; cystic fibrosis transmembrane conductance regulator; 5T; 7T;  
KW 9T; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200121833-A2.  
XX  
PD 29-MAR-2001.

XX  
PF 19-SEP-2000; 2000WO-GB03597.  
XX  
PR 24-SEP-1999; 99GB-0022527.  
XX  
PA (ASTR ) ASTRAZENECA AB.  
PA (ASTR ) ASTRAZENECA UK LTD.  
XX  
PI Bayliffe AI, Doctor E, Kelly SJ, Robertson NH;  
XX  
DR WPI; 2001-308093/32.

XX  
PT Detecting 5T, 7T and 9T alleles in intron 8 of the human cystic  
PT fibrosis transmembrane conductance regulator comprises contacting a  
PT nucleic acid sample with a multiplex of diagnostic primers comprising  
PT 5T, 7T and 9T variant primers -  
XX  
PS Disclosure; Page 6; 73pp; English.

XX  
CC The present invention relates to detecting 5T, 7T and 9T alleles in  
CC intron 8 of the human cystic fibrosis transmembrane conductance  
CC regulator (CFTR) gene. The method involves a multiplex of  
CC diagnostic primers comprising 5T, 7T and 9T variant primers.  
CC The invention is useful for simultaneous detection of variable length  
CC polythymidine (polyT) tract alleles in the CFTR gene that are  
CC associated with the phenotypic modulation of selected CFTR mutations.  
CC The present sequence is a primer used in the invention.

XX  
SQ Sequence 11 BP; 8 A; 1 C; 0 G; 1 T; 1 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAH19036 ..

Align seg 1/1 to: AAH19036 from: 1 to: 11

1 LysLysAsn 3  
|||||  
2 AAAAAAAT 10

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF29873

seq\_documentation\_block:  
ID AAF29873 standard; DNA; 11 BP.

XX  
AC AAF29873;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Starting 1 fragment #6.  
XX  
KW Cloning; exon shuffling; store; adapter; ss.  
XX  
OS Unidentified.  
XX  
PN WO200100816-A1.  
XX  
PD 04-JAN-2001.  
XX  
PF 27-JUN-2000; 2000WO-GB02512.  
XX  
PR 28-JUN-1999; 99NO-0001325.  
PR 20-JUN-2000; 2000NO-0003190.  
PR 20-JUN-2000; 2000NO-0003191.  
XX  
PA (COMP-) COMPLETE GENOMICS AS.  
PA (JONE/) JONES E L.  
XX  
PI Lexow P;  
XX  
DR WPI; 2001-123006/13.

XX  
PT Attaching fragments of first nucleic acids to second nucleic acids by  
PT use of adapters complementary to first single stranded regions on the  
PT first molecules but which have a different single stranded region at  
PT the other terminus -  
XX  
PS Disclosure; Fig 2; 100pp; English.  
XX  
CC The present invention relates to attaching a fragment of first and  
CC second nucleic acid molecules involves use of an adapter molecule  
CC which is complementary to a single stranded region generated on the  
CC target but which has a different single stranded region at its  
CC other terminus and therefore modifies single stranded regions  
CC presented for binding by the target. Attaching first and second  
CC nucleic acid molecules may be used in cloning. The method can  
CC also be used for exon shuffling other recombinations that are  
CC relevant in connection with artificial evolutionary systems.  
CC The fragment chains may be used to store information.

XX  
SQ Sequence 11 BP; 0 A; 1 C; 2 G; 8 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAF29873/rev ..

Align seg 1/1 to reverse of: AAF29873 from: 1 to: 11

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3



seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT73560

seq\_documentation\_block:

ID AAT73560 standard; RNA; 12 BP.

XX AC AAT73560;

XX DT 29-AUG-1997 (first entry)

XX DE Antiviral capped short oligonucleotide 2.

XX KW antiviral; diphosphate; triphosphate; capped; guanylyl transferase;  
XX large scale synthesis; influenza RNA polymerase; inhibit; cap-dependent;  
XX transcription; in vitro; study; transcription; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT modified\_base 1

FT /\*tag= a

FT /note= "m7G"

FT modified\_base 2

FT /\*tag= b

FT /note= "pppU"

XX WO9641812-A1.

XX PD 27-DEC-1996.

XX PF 07-JUN-1996; 96WO-GB01360.

XX PR 09-JUN-1995; 95GB-0011720.

XX PA (ISIS-) ISIS INNOVATION LTD.

XX PI Brownlee GG, Fodor E, Lowe G;

XX DR WPI; 1997-065418/06.

XX PT Oligo:nucleotide phosphorylation process - for prepn. of  
XX oligo:nucleotide(s) having 5'-terminal di- or triphosphate gp.,  
XX which are useful in influenza virus transcription studies

XX PS Claim 11; Page 26; 35pp; English.

XX CC A novel process comprises reacting an oligonucleotide, having a 5'  
XX -terminal -OH gp., with a phosphorylating or phosphorylating agent,  
XX then with a phosphate or pyrophosphate and recovering an  
XX oligonucleotide, having a 5'-terminal di- or triphosphate group. The  
XX recovered oligonucleotides (AAT73560-63) are then capped by reaction with  
XX guanylyl transferase. The process is useful for the large scale synthesis  
XX of short-capped oligonucleotides. The short-capped oligonucleotides are  
XX potential antiviral agents as they can bind to influenza RNA polymerase,  
XX and specifically inhibit cap-dependent transcription in vitro. The  
XX prods. may also be useful for the study of influenza virus  
XX transcription.

SQ Sequence 12 BP; 0 A; 0 C; 1 G; 11 U; 0 other;

alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AAT73560/rev ..

Align seg 1/1 to reverse of: AAT73560 from: 1 to: 12

1 LysLysAsn 3

|||||||

9 AAAAAAAC 1

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAA94316

seq\_documentation\_block:

ID AAA94316 standard; DNA; 12 BP.

XX AC AAA94316;

XX DT 11-JAN-2001 (first entry)

XX DE RNA-protein fusion oligonucleotide 13-P.

XX KW RNA-protein fusion; protein library; protein isolation; gene cloning; ss.  
XX OS Synthetic.

XX FH Key Location/Qualifiers

FT modified\_base 12

FT /\*tag= a

FT /mod\_base= OTHER

FT /note= "attached to puromycin, a peptide acceptor"

XX PN WO200047775-A1.

XX PD 17-AUG-2000.

XX PF 01-FEB-2000; 2000WO-US02589.

XX PR 09-FEB-1999; 99US-0247190.

XX PA (GEO ) GEN HOSPITAL CORP.

XX PI Szostak JW, Roberts RW, Liu R;

XX DR WPI; 2000-533022/48.

XX PT Producing protein or DNA libraries which are useful for improving  
XX existing proteins, by in vitro translating protein coding sequences to  
XX produce RNA-protein fusions and incubating these protein fusions under  
XX high salt conditions

XX PS Disclosure; Page 43; 121pp; English.

XX CC The present sequence is one of a number of oligonucleotides which were  
XX used for the generation of RNA-protein fusions, including fusions having  
XX a myc epitope tag. The RNA-protein fusions comprise a protein covalently  
XX linked to the 3' end of its own mRNA. This is accomplished by synthesis  
XX and in vitro or in situ translation of an mRNA molecule with a peptide  
XX acceptor attached to its 3' end. The RNA-protein fusions are incubated  
XX under high salt conditions to produce a protein library. This method is  
XX useful for improving or altering existing proteins, as well as for  
XX isolating new proteins and nucleic acid or small molecule targets. It may  
XX also be used to improve human or humanised single-chain antibodies for  
XX the treatment of a number of diseases. The method is useful for the  
XX isolation of proteins with specific binding properties, for screening  
XX cDNA libraries and cloning new genes on the basis of protein-protein  
XX interactions. Unlike prior art, the new method does not rely on  
XX maintaining the integrity of an mRNA:ribosome:nascent chain ternary  
XX complex, which is very fragile and is therefore of limited use. The  
XX method does not rely on topological links between the protein and the  
XX nucleic acid so that the information of the protein is retained and can  
XX be recovered in readable, nucleic acid form.

SQ Sequence 12 BP; 10 A; 2 C; 0 G; 0 U; 0 other;

alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AAA94316 ..



Align seg 1/1 to: AAA94316 from: 1 to: 12

1 LysLysAsn 3  
|||||  
3 AAAAAAAC 11

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAA10356

seq\_documentation\_block:

ID AAA10356 standard; DNA; 12 BP.

AC AAA10356;

DT 03-JUL-2000 (first entry)

DE DNA ligand binding assay competitor oligonucleotide, SEQ ID NO:39.

KW Nucleic acid ligand binding assay; duplex formation; stability;

KW detectable signal; competition assay; competitor oligonucleotide; ds.

OS Synthetic.

PN WO200015848-A1.

PD 23-MAR-2000.

PF 10-SEP-1999; 99WO-US20719.

PR 11-SEP-1998; 98US-0151890.

PA (GENE-) GENELABS TECHNOLOGIES INC.

PI Schroth GP, Bruice TW, Suh YJ;

DR WPT; 2000-271478/23.

PT Determining binding affinity of a ligand to an oligonucleotide sequence  
PT in double stranded form, comprises measuring the effect of adding  
PT increasing amounts of a ligand on a signal generated by two indicator  
PT oligonucleotides of the duplex -

PS Examples 7-8; Page 20; 78pp; English.

XX The invention relates to new methods of determining the binding affinity  
CC of a ligand to an oligonucleotide sequence, particularly to a duplex.  
CC The ligand is typically a metal ion, a small organic or inorganic  
CC molecule, a protein or a multi-protein complex. The methods comprise  
CC measuring the effect of adding increasing amounts of a ligand on a signal  
CC generated by two indicator oligonucleotides of the duplex. In the absence  
CC of ligand, conditions are such that the oligonucleotides exist primarily  
CC in single-stranded form; binding of ligand to double-stranded nucleic  
CC acids stabilises the duplexes, such that duplex formation is favoured.  
CC One of the indicator oligonucleotides contains a first group capable of  
CC producing a detectable signal, while the other indicator oligonucleotide  
CC contains a second group that on hybridisation of the two indicator  
CC molecules, will detectably alter the signal produced by the first group.  
CC The signal may be increased or decreased on hybridisation. For example,  
CC the pairs of signalling groups used could be a radioactive group and a  
CC scintillant (where an increase in signal intensity indicates that  
CC hybridisation has taken place) or a fluorophore and a fluorescence  
CC quencher (where a reduction in signal intensity indicates that  
CC hybridisation has occurred). Other methods of the invention comprise a  
CC strand displacement assay, where the ability of an unlabelled  
CC is determined in the absence and presence of ligand; and a competition  
CC assay, where an unlabelled single or double-stranded competitor  
CC oligonucleotide is added to the ligand-bound indicator duplex, and the  
CC effect on the signal produced from the indicator duplex determined. The  
CC methods are useful for determining the binding affinity of a ligand to  
CC an oligonucleotide sequence. They are particularly useful for  
CC determining relative binding affinities of various ligands to various  
CC oligonucleotide sequences, particularly double-stranded oligonucleotide

CC sequences. The assays allow rapid and convenient determination of nucleic  
CC acid binding specificities. Sequences AAA10342-A10391 represent  
CC competitor oligonucleotides used in competition assays in  
CC exemplifications of the present invention.

XX  
SQ Sequence 12 BP; 8 A; 4 C; 0 G; 0 U; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AAA10356 ..

Align seg 1/1 to: AAA10356 from: 1 to: 12

1 LysLysAsn 3

|||||

3 AAAAAAAC 11

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH46026

seq\_documentation\_block:

ID AAH46026 standard; DNA; 12 BP.

AC AAH46026;

DT 12-SEP-2001 (first entry)

DE Mycobacterium phlei murA oligonucleotide 3.

KW Synthetic oligonucleotide; dinucleotide repeat; cytostatic; apoptosis;

KW cell cycle arrest; cell proliferation; caspase; cytokine; interleukin;

KW tumour necrosis factor; TNF; cancer; carcinoma; sarcoma; leukemia;

KW lymphoma; ss; murA.

OS Mycobacterium phlei.

OS Synthetic.

PN WO200144465-A2.

PD 21-JUN-2001.

PF 12-DEC-2000; 2000WO-CA01467.

PR 13-DEC-1999; 99US-0170325.

PR 29-AUG-2000; 2000US-0228925.

PA (BION-) BIONICHE LIFE SCI INC.

PI Phillips NC, Fillion MC;

DR WPT; 2001-398150/42.

PT Composition comprising synthetic oligonucleotides which comprise  
PT multiple repeats of dinucleotides such as GT, TG useful for treating  
PT cancer by inducing cell cycle arrest, inhibiting proliferation,  
PT activating caspases -

PS Claim 5; Page 41; 77pp; English.

XX The present sequence is that of a synthetic Mycobacterium phlei murA  
CC derived oligonucleotide useful to the invention. The invention relates to  
CC a composition, comprising a 2 to 20 base 3'-OH, 5'-OH synthetic  
CC oligonucleotide which comprises multiple repeats of dinucleotides such as  
CC GT, TG, etc., according to specific formula and having cytostatic  
CC activity. The oligonucleotide compositions are useful for inducing cell  
CC cycle arrest, inhibition of proliferation, activation of caspases and  
CC induction of apoptosis or production of cytokines such as interleukin  
CC (IL)-1-beta, IL-6, IL-10, IL-12 and tumour necrosis factor (TNF)-alpha by  
CC immune system cells, in an animal having cancer such as primary



CC carcinoma, secondary carcinoma, primary sarcoma and secondary sarcoma  
CC such as, leukemia, lymphoma, breast, prostate, colorectal, ovarian or  
CC bone cancer. The compositions induce apoptosis independent of Fas,  
CC p53/p21, p21/waf-1/CIP, p15(ink4B), p16(ink4), drug resistance,  
CC caspase 3, transforming growth factor (TGF)-beta 1 receptor and hormone  
CC dependence.  
XX  
SQ Sequence 12 BP; 0 A; 0 C; 2 G; 10 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AAH46026/rev ..

Align seg 1/1 to reverse of: AAH46026 from: 1 to: 12

1 LysLysAsn 3  
|||||  
11 AAAAAAA 3

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABH67135

seq\_documentation\_block:

ID\_ ABH67135 standard; DNA; 12 BP.

XX AC ABH67135;  
XX DT 22-FEB-2002 (first entry)  
XX DE Oligonucleotide primer SEQ ID NO 267112 for detecting SNP TSC00000016.  
XX  
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
XX Homo sapiens.

XX WO200177384-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-IB00713.

XX 07-APR-2000; 2000DE-1019173.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
XX designed to detect single nucleotide polymorphisms and cytosine  
XX methylation status

XX Claim 1; SEQ ID 267112; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
XX acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
XX and cytosine methylation status in chemically pretreated genomic DNA. The  
XX oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
XX range of diseases including immune system, gastrointestinal, respiratory,  
XX central nervous system, cardiovascular and metabolic disorders. The  
XX oligomers are also used for detecting cell type differentiation.

XX ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

XX ABT00010-ABT82073 represent the oligomers described in the invention.

XX NOTE: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 12 BP; 2 A; 0 C; 0 G; 10 T; 0 other;  
  
alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x ABH67135/rev ..

Align seg 1/1 to reverse of: ABH67135 from: 1 to: 12

1 LysLysAsn 3  
|||||  
9 AAAAAAA 1

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABH67157

seq\_documentation\_block:

ID\_ ABH67157 standard; DNA; 12 BP.

XX AC ABH67157;  
XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 267134 for detecting SNP TSC00000024.  
XX  
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
XX Homo sapiens.

XX WO200177384-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-IB00713.

XX 07-APR-2000; 2000DE-1019173.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
XX designed to detect single nucleotide polymorphisms and cytosine  
XX methylation status

XX Claim 1; SEQ ID 267134; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
XX acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
XX and cytosine methylation status in chemically pretreated genomic DNA. The  
XX oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
XX range of diseases including immune system, gastrointestinal, respiratory,  
XX central nervous system, cardiovascular and metabolic disorders. The  
XX oligomers are also used for detecting cell type differentiation.

XX ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

XX ABT00010-ABT82073 represent the oligomers described in the invention.

XX NOTE: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12 BP; 1 A; 0 C; 2 G; 9 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3



Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH67157/rev ..

Align seg 1/1 to reverse of: ABH67157 from: 1 to: 12

1 LysLysAsn 3  
|||||  
10 AAAAAAAC 4

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67218

## seq\_documentation\_block:

ID ABH67218 standard; DNA; 12 BP.

XX

AC ABH67218;

DT 22-FEB-2002 (first entry)

DE Oligonucleotide primer SEQ ID NO 267195 for detecting SNP TSC00000048.

KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

OS Homo sapiens.

PN WO200177384-A2.

XX

PD 18-OCT-2001.

PF 06-APR-2001; 2001WO-IB00713.

XX

PR 07-APR-2000; 2000DE-1019173.

XX

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX

DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is

PT designed to detect single nucleotide polymorphisms and cytosine

PT methylation status -

XX Claim 1; SEQ ID 267195; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic

CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)

CC and cytosine methylation status in chemically pretreated genomic DNA. The

CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a

CC range of diseases including immune system, gastrointestinal, respiratory,

CC central nervous system, cardiovascular and metabolic disorders. The

CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABT00010-ABT2073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12 BP; 2 A; 0 C; 0 G; 10 T; 0 other;

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH67218/rev ..

Align seg 1/1 to reverse of: ABH67218 from: 1 to: 12

1 LysLysAsn 3  
|||||  
10 AAAAAAAC 2

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67278

## seq\_documentation\_block:

ID ABH67278 standard; DNA; 12 BP.

XX

AC ABH67278;

DT 22-FEB-2002 (first entry)

DE Oligonucleotide primer SEQ ID NO 267255 for detecting SNP TSC00000067.

KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

OS Homo sapiens.

PN WO200177384-A2.

XX

PD 18-OCT-2001.

PF 06-APR-2001; 2001WO-IB00713.

XX

PR 07-APR-2000; 2000DE-1019173.

XX

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX

DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is

PT designed to detect single nucleotide polymorphisms and cytosine

PT methylation status -

XX Claim 1; SEQ ID 267255; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic

CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)

CC and cytosine methylation status in chemically pretreated genomic DNA. The

CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a

CC range of diseases including immune system, gastrointestinal, respiratory,

CC central nervous system, cardiovascular and metabolic disorders. The

CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABT00010-ABT2073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12 BP; 9 A; 1 C; 0 G; 2 T; 0 other;

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH67278 ..

Align seg 1/1 to: ABH67278 from: 1 to: 12

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10



seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABH67354

seq\_documentation\_block:

ID ABH67354 standard; DNA; 12 BP.

XX AC ABH67354;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 267331 for detecting SNP TSC0000112.

XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN W0200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPiG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status

XX PS Claim 1; SEQ ID 267331; 29pp + Sequence Listing; German.

XX CC This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABI00010-ABI82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12 BP; 0 A; 0 C; 4 G; 8 T; 0 other;

alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67354/rev ..

Align seg 1/1 to reverse of: ABH67354 from: 1 to: 12

1 LysLysAsn 3

|||||

12 AAAAAAAC 4

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABH67386

seq\_documentation\_block:

ID ABH67386 standard; DNA; 12 BP.

XX AC ABH67386;

XX DT

XX DE Oligonucleotide primer SEQ ID NO 267363 for detecting SNP TSC0000141.

XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN W0200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPiG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status

XX PS Claim 1; SEQ ID 267363; 29pp + Sequence Listing; German.

XX CC This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABI00010-ABI82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12 BP; 10 A; 0 C; 0 G; 2 T; 0 other;

alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67386 ..

Align seg 1/1 to: ABH67386 from: 1 to: 12

1 LysLysAsn 3

|||||

4 AAAAAAAT 12

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABH67427

seq\_documentation\_block:

ID ABH67427 standard; DNA; 12 BP.

XX AC ABH67427;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 267404 for detecting SNP TSC0000178.

XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;







```

XX PA (EPiG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX PT designed to detect single nucleotide polymorphisms and cytosine
XX PT methylation status
XX
XX Claim 1; SEQ ID 267459; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
XX CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX CC and cytosine methylation status in chemically pretreated genomic DNA. The
XX CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX CC range of diseases including immune system, gastrointestinal, respiratory,
XX CC central nervous system, cardiovascular and metabolic disorders. The
XX CC oligomers are also used for detecting cell type differentiation.
XX CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX CC ABT00010-ABT82073 represent the oligomers described in the invention.
XX CC NOTE: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 12 BP; 1 A; 0 C; 3 G; 8 T; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH67482/rev ..
    1 LysLysAsn 3
    |||||
    9 AAAAAAAC 1

Align seg 1/1 to reverse of: ABH67482 from: 1 to: 12

seq_name: /SIDS5/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:ABH67505

seq_documentation_block:
ID_ ABH67505 standard; DNA; 12 BP.
XX
XX AC ABH67505;
XX
XX 22-FEB-2002 (first entry)
XX
XX Oligonucleotide primer SEQ ID NO 267482 for detecting SNP TSC0000294.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX PT designed to detect single nucleotide polymorphisms and cytosine
XX PT methylation status
XX
XX Claim 1; SEQ ID 267482; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
XX CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX CC and cytosine methylation status in chemically pretreated genomic DNA. The
XX CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX CC range of diseases including immune system, gastrointestinal, respiratory,
XX CC central nervous system, cardiovascular and metabolic disorders. The
XX CC oligomers are also used for detecting cell type differentiation.
XX CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX CC ABT00010-ABT82073 represent the oligomers described in the invention.
XX CC NOTE: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 12 BP; 1 A; 0 C; 3 G; 8 T; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH67482/rev ..
    1 LysLysAsn 3
    |||||
    9 AAAAAAAC 1

Align seg 1/1 to reverse of: ABH67482 from: 1 to: 12

seq_name: /SIDS5/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:ABH67505

seq_documentation_block:
ID_ ABH67505 standard; DNA; 12 BP.
XX
XX AC ABH67505;
XX
XX 22-FEB-2002 (first entry)
XX
XX Oligonucleotide primer SEQ ID NO 267482 for detecting SNP TSC0000400.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX PT designed to detect single nucleotide polymorphisms and cytosine
XX PT methylation status
XX
XX Claim 1; SEQ ID 267617; 29pp + Sequence Listing; German.

```



XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 12 BP; 1 A; 0 C; 1 G; 10 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67640/rev ..

Align seg 1/1 to reverse of: ABH67640 from: 1 to: 12

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.ABH67643

seq\_documentation\_block:

ID ABH67643 standard; DNA; 12 BP.

XX AC ABH67643;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 267620 for detecting SNP TSC0000401.

XX SNF: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX

OS Homo sapiens.

XX PN WO200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status -

PS Claim 1; SEQ ID 267620; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC

CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 12 BP; 10 A; 1 C; 0 G; 1 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67643 ..

Align seg 1/1 to: ABH67643 from: 1 to: 12

1 LysLysAsn 3  
|||||  
4 AAAAAAAC 12

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.ABH67751

seq\_documentation\_block:

ID ABH67751 standard; DNA; 12 BP.

XX AC ABH67751;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 267728 for detecting SNP TSC0000491.

XX SNF: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX

OS Homo sapiens.

XX PN WO200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status -

PS Claim 1; SEQ ID 267728; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at



```
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 12 BP; 2 A; 0 C; 1 G; 9 T; 0 other;
SQ

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x ABH67751/rev ..
  Align seg 1/1 to reverse of: ABH67751 from: 1 to: 12
  1 LysLysAsn 3
  |||||
  10 AAAAAAAC 2

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67810
seq_documentation_block:
  ID ABH67810 standard; DNA; 12 BP.
  XX
  AC ABH67810;
  DT 22-FEB-2002 (first entry)
  XX
  DE Oligonucleotide primer SEQ ID NO 267787 for detecting SNP TSC0000529.
  XX
  KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
  KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
  KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
  XX
  OS Homo sapiens.
  XX
  PN WO200177384-A2.
  PD 18-OCT-2001.
  XX
  PF 06-APR-2001; 2001WO-IB00713.
  XX
  PR 07-APR-2000; 2000DE-1019173.
  XX
  PA (EPIG-) EPIGENOMICS AG.
  XX
  PI Olek A, Piepenbrock C, Berlin K;
  XX
  DR WPI; 2001-657177/75.
  XX
  PT Set of oligonucleotides, useful for diagnosis and cell typing, is
  PT designed to detect single nucleotide polymorphisms and cytosine
  PT methylation status
  XX
  PS Claim 1; SEQ ID 267787; 29pp + Sequence Listing; German.
  XX
  CC This invention describes novel oligonucleotide primers or peptide nucleic
  CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
  CC and cytosine methylation status in chemically pretreated genomic DNA. The
  CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
  CC range of diseases including immune system, gastrointestinal, respiratory,
  CC central nervous system, cardiovascular and metabolic disorders. The
  CC oligomers are also used for detecting cell type differentiation.
  CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
  CC ABI00010-ABI82073 represent the oligomers described in the invention.
  CC NOTE: The sequence data for this patent did not form part of the printed
  CC specification, but was obtained in electronic format from WIPO at
  CC ftp.wipo.int/pub/published_pct_sequences.
  XX
  SQ Sequence 12 BP; 10 A; 1 C; 0 G; 1 T; 0 other;

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x ABH67912 ..
  seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67912
  seq_documentation_block:
  ID ABH67912 standard; DNA; 12 BP.
  XX
  AC ABH67912;
  DT 22-FEB-2002 (first entry)
  XX
  DE Oligonucleotide primer SEQ ID NO 267889 for detecting SNP TSC0000635.
  XX
  KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
  KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
  KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
  XX
  OS Homo sapiens.
  XX
  PN WO200177384-A2.
  PD 18-OCT-2001.
  XX
  PF 06-APR-2001; 2001WO-IB00713.
  XX
  PR 07-APR-2000; 2000DE-1019173.
  XX
  PA (EPIG-) EPIGENOMICS AG.
  XX
  PI Olek A, Piepenbrock C, Berlin K;
  XX
  DR WPI; 2001-657177/75.
  XX
  PT Set of oligonucleotides, useful for diagnosis and cell typing, is
  PT designed to detect single nucleotide polymorphisms and cytosine
  PT methylation status
  XX
  PS Claim 1; SEQ ID 267889; 29pp + Sequence Listing; German.
  XX
  CC This invention describes novel oligonucleotide primers or peptide nucleic
  CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
  CC and cytosine methylation status in chemically pretreated genomic DNA. The
  CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
  CC range of diseases including immune system, gastrointestinal, respiratory,
  CC central nervous system, cardiovascular and metabolic disorders. The
  CC oligomers are also used for detecting cell type differentiation.
  CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
  CC ABI00010-ABI82073 represent the oligomers described in the invention.
  CC NOTE: The sequence data for this patent did not form part of the printed
  CC specification, but was obtained in electronic format from WIPO at
  CC ftp.wipo.int/pub/published_pct_sequences.
  XX
  SQ Sequence 12 BP; 9 A; 0 C; 1 G; 2 T; 0 other;

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x ABH67912 ..
  seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67912
  seq_documentation_block:
  ID ABH67912 standard; DNA; 12 BP.
  XX
  AC ABH67912;
  DT 22-FEB-2002 (first entry)
  XX
  DE Oligonucleotide primer SEQ ID NO 267889 for detecting SNP TSC0000635.
  XX
  KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
  KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
  KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
  XX
  OS Homo sapiens.
  XX
  PN WO200177384-A2.
  PD 18-OCT-2001.
  XX
  PF 06-APR-2001; 2001WO-IB00713.
  XX
  PR 07-APR-2000; 2000DE-1019173.
  XX
  PA (EPIG-) EPIGENOMICS AG.
  XX
  PI Olek A, Piepenbrock C, Berlin K;
  XX
  DR WPI; 2001-657177/75.
  XX
  PT Set of oligonucleotides, useful for diagnosis and cell typing, is
  PT designed to detect single nucleotide polymorphisms and cytosine
  PT methylation status
  XX
  PS Claim 1; SEQ ID 267889; 29pp + Sequence Listing; German.
  XX
  CC This invention describes novel oligonucleotide primers or peptide nucleic
  CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
  CC and cytosine methylation status in chemically pretreated genomic DNA. The
  CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
  CC range of diseases including immune system, gastrointestinal, respiratory,
  CC central nervous system, cardiovascular and metabolic disorders. The
  CC oligomers are also used for detecting cell type differentiation.
  CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
  CC ABI00010-ABI82073 represent the oligomers described in the invention.
  CC NOTE: The sequence data for this patent did not form part of the printed
  CC specification, but was obtained in electronic format from WIPO at
  CC ftp.wipo.int/pub/published_pct_sequences.
  XX
  SQ Sequence 12 BP; 9 A; 0 C; 1 G; 2 T; 0 other;
```



Align seg 1/1 to: ABH67912 from: 1 to: 12

1 LysLysAsn 3  
|||||  
1 AAAAAAAT 9

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67913

seq\_documentation\_block:

ID ABH67913 standard; DNA; 12 BP.

XX

AC ABH67913;

XX

DT 22-FEB-2002 (first entry)

XX

DE Oligonucleotide primer SEQ ID NO 267890 for detecting SNP TSC0000635.

XX

SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX

OS Homo sapiens.

XX

PN WO200177384-A2.

XX

PD 18-OCT-2001.

XX

PF 06-APR-2001; 2001WO-IB00713.

XX

PR 07-APR-2000; 2000DE-1019173.

XX

PA (EPIC-) EPIGENOMICS AG.

XX

PI Olek A, Piepenbrock C, Berlin K;

XX

DR WPI; 2001-657177/75.

XX

Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine

PT methylation status -

XX

PS Claim 1; SEQ ID 267890; 29pp + Sequence Listing; German.

XX

This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABI00010-ABI82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 12 BP; 10 A; 0 C; 0 G; 2 T; 0 other;

alignment\_scores:

Quality: 16.00

Ratio: 5.333

Percent Similarity: 100.000

Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67913 ..

Align seg 1/1 to: ABH67913 from: 1 to: 12

1 LysLysAsn 3  
|||||  
1 AAAAAAAT 9

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH67928

seq\_documentation\_block:

ID ABH67928 standard; DNA; 12 BP.

XX

AC ABH67928;

XX

DT 22-FEB-2002 (first entry)

XX

DE Oligonucleotide primer SEQ ID NO 267905 for detecting SNP TSC0000655.

XX

SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX

OS Homo sapiens.

XX

PN WO200177384-A2.

XX

PD 18-OCT-2001.

XX

PF 06-APR-2001; 2001WO-IB00713.

XX

PR 07-APR-2000; 2000DE-1019173.

XX

PA (EPIC-) EPIGENOMICS AG.

XX

PI Olek A, Piepenbrock C, Berlin K;

XX

DR WPI; 2001-657177/75.

XX

Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine

PT methylation status -

XX

PS Claim 1; SEQ ID 267905; 29pp + Sequence Listing; German.

XX

This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.

CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

CC ABI00010-ABI82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 12 BP; 1 A; 0 C; 2 G; 9 T; 0 other;

alignment\_scores:

Quality: 16.00

Ratio: 5.333

Percent Similarity: 100.000

Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH67928/rev ..

Align seg 1/1 to reverse of: ABH67928 from: 1 to: 12

1 LysLysAsn 3  
|||||  
9 AAAAAAAT 1

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68005

seq\_documentation\_block:

ID ABH68005 standard; DNA; 12 BP.

XX



AC ABH68005;  
XX  
DT 22-FEB-2002 (first entry)  
XX  
DE Oligonucleotide primer SEQ ID NO 267982 for detecting SNP TSC00000755.  
XX  
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
OS Homo sapiens.  
XX  
PN W0200177384-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 06-APR-2001; 2001WO-IB00713.  
XX  
PR 07-APR-2000; 2000DE-1019173.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2001-657177/75.  
XX  
XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status  
XX  
PS Claim 1; SEQ ID 267982; 29pp + Sequence Listing; German.  
XX  
XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC AB100010-AB182073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 12 BP; 2 A; 0 C; 0 G; 10 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x ABH68005/rev ..  
Align seg 1/1 to reverse of: ABH68005 from: 1 to: 12

1 LysLysAsn 3  
|||||  
10 AAAAAAAT 2

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68036  
seq\_documentation\_block:  
ID ABH68036 standard; DNA; 12 BP.  
XX  
AC ABH68036;  
XX  
DT 22-FEB-2002 (first entry)  
XX  
DE Oligonucleotide primer SEQ ID NO 268013 for detecting SNP TSC00000799.  
XX

KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
OS Homo sapiens.  
XX  
PN W0200177384-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 06-APR-2001; 2001WO-IB00713.  
XX  
PR 07-APR-2000; 2000DE-1019173.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2001-657177/75.  
XX  
XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status  
XX  
PS Claim 1; SEQ ID 268013; 29pp + Sequence Listing; German.  
XX  
XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC AB100010-AB182073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 12 BP; 9 A; 2 C; 0 G; 1 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x ABH68036 ..  
Align seg 1/1 to: ABH68036 from: 1 to: 12

1 LysLysAsn 3  
|||||  
2 AAAAAAAC 10

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68090  
seq\_documentation\_block:  
ID ABH68090 standard; DNA; 12 BP.  
XX  
AC ABH68090;  
XX  
DT 22-FEB-2002 (first entry)  
XX  
DE Oligonucleotide primer SEQ ID NO 268067 for detecting SNP TSC00000855.  
XX  
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
OS Homo sapiens.  
XX



```

PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
XX PF 06-APR-2001; 2001WO-IB00713.
XX
XX PR 07-APR-2000; 2000DE-1019173.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX
XX DR WPI; 2001-657177/75.
XX
XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is
XX PT designed to detect single nucleotide polymorphisms and cytosine
XX PT methylation status -
XX
XX PS Claim 1; SEQ ID 268067; 29pp + Sequence Listing; German.
XX
XX CC This invention describes novel oligonucleotide primers or peptide nucleic
XX CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX CC and cytosine methylation status in chemically pretreated genomic DNA. The
XX CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX CC range of diseases including immune system, gastrointestinal, respiratory,
XX CC central nervous system, cardiovascular and metabolic disorders. The
XX CC oligomers are also used for detecting cell type differentiation.
XX CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX CC ABI00010-ABI82073 represent the oligomers described in the invention.
XX CC NOTE: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 12 BP; 2 A; 0 C; 1 G; 9 T; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH68090/rev ..

Align seg 1/1 to reverse of: ABH68090 from: 1 to: 12

1 LysLysAsn 3
|||||
9 AAAAAAAAAAC 1

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68175

seq_documentation_block:
ID ABH68175 standard; DNA; 12 BP.
XX
XX AC ABH68175;
XX
XX DT 22-FEB-2002 (first entry)
XX
XX DE Oligonucleotide primer SEQ ID NO 268152 for detecting SNP TSC0000930.
XX
XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX OS Homo sapiens.
XX
XX PN WO200177384-A2.
XX
XX PD 18-OCT-2001.
XX
XX PF 06-APR-2001; 2001WO-IB00713.
XX

```

```

PR 07-APR-2000; 2000DE-1019173.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX
XX DR WPI; 2001-657177/75.
XX
XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is
XX PT designed to detect single nucleotide polymorphisms and cytosine
XX PT methylation status -
XX
XX PS Claim 1; SEQ ID 268152; 29pp + Sequence Listing; German.
XX
XX CC This invention describes novel oligonucleotide primers or peptide nucleic
XX CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX CC and cytosine methylation status in chemically pretreated genomic DNA. The
XX CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX CC range of diseases including immune system, gastrointestinal, respiratory,
XX CC central nervous system, cardiovascular and metabolic disorders. The
XX CC oligomers are also used for detecting cell type differentiation.
XX CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX CC ABI00010-ABI82073 represent the oligomers described in the invention.
XX CC NOTE: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 12 BP; 9 A; 2 C; 0 G; 1 T; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH68175 ..

Align seg 1/1 to: ABH68175 from: 1 to: 12

1 LysLysAsn 3
|||||
1 AAAAAAAAAAT 9

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68195

seq_documentation_block:
ID ABH68195 standard; DNA; 12 BP.
XX
XX AC ABH68195;
XX
XX DT 22-FEB-2002 (first entry)
XX
XX DE Oligonucleotide primer SEQ ID NO 268172 for detecting SNP TSC0000936.
XX
XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX OS Homo sapiens.
XX
XX PN WO200177384-A2.
XX
XX PD 18-OCT-2001.
XX
XX PF 06-APR-2001; 2001WO-IB00713.
XX
XX PR 07-APR-2000; 2000DE-1019173.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX

```



```
DR WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
XX
PS Claim 1; SEQ ID 268172; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABT00010-ABT82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 11 A; 1 C; 0 G; 0 U; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH68195 ..
    1 LysLysAsn 3
    |||||
    12 AAAAAAAC 4

Align seg 1/1 to: ABH68195 from: 1 to: 12

seq_name: /SIDS5/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68196
seq_documentation_block:
ID ABH68196 standard; DNA; 12 BP.
XX
XX
AC ABH68196;
XX
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 268173 for detecting SNP TSC00000936.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
OS
XX
XX WO200177384-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-IB00713.
PF
XX
XX 07-APR-2000; 2000DE-1019173.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2001-657177/75.
DR
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
XX

DR WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
XX
PS Claim 1; SEQ ID 268173; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABT00010-ABT82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 11 A; 1 C; 0 G; 0 U; 0 other;

alignment_scores:
    Quality: 16.00      Length: 3
    Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x ABH68195 ..
    1 LysLysAsn 3
    |||||
    12 AAAAAAAC 9

Align seg 1/1 to: ABH68195 from: 1 to: 12

seq_name: /SIDS5/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68196
seq_documentation_block:
ID ABH68196 standard; DNA; 12 BP.
XX
XX
AC ABH68196;
XX
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 268173 for detecting SNP TSC00000936.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
OS
XX
XX WO200177384-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-IB00713.
PF
XX
XX 07-APR-2000; 2000DE-1019173.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2001-657177/75.
DR
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
XX
```



CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 12 BP; 2 A; 0 C; 0 G; 10 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH68217/rev ..

Align seg 1/1 to reverse of: ABH68217 from: 1 to: 12

1 LysLysAsn 3  
|||||||  
12 AAAAAAAAAAT 4

seq\_name: /SID55/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68243

seq\_documentation\_block:

ID ABH68243 standard; DNA; 12 BP.

XX AC ABH68243;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 268220 for detecting SNP TSC0000988.

XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN WO200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status -

XX Claim 1; SEQ ID 268220; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 12 BP; 1 A; 2 C; 0 G; 9 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x ABH68243/rev ..

Align seg 1/1 to reverse of: ABH68243 from: 1 to: 12

1 LysLysAsn 3  
|||||||  
11 AAGAAAAAT 3

seq\_name: /SID55/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:ABH68246

seq\_documentation\_block:

ID ABH68246 standard; DNA; 12 BP.

XX AC ABH68246;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 268223 for detecting SNP TSC0000988.

XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN WO200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is  
PT designed to detect single nucleotide polymorphisms and cytosine  
PT methylation status -

XX Claim 1; SEQ ID 268223; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic  
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)  
CC and cytosine methylation status in chemically pretreated genomic DNA. The  
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a  
CC range of diseases including immune system, gastrointestinal, respiratory,  
CC central nervous system, cardiovascular and metabolic disorders. The  
CC oligomers are also used for detecting cell type differentiation.  
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and  
CC ABT00010-ABT82073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 12 BP; 1 A; 0 C; 0 G; 11 T; 0 other;



alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH68246/rev ..

Align seg 1/1 to reverse of: ABH68246 from: 1 to: 12

1 LysLysAsn 3  
|||||  
11 AAAAAAAT 3

seq\_name: /SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68268

## seq\_documentation\_block:

ID ABH68268 standard; DNA; 12 BP.

XX AC ABH68268;  
XX DT 22-FEB-2002 (first entry)  
XX DE Oligonucleotide primer SEQ ID NO 268245 for detecting SNP TSC0001000.  
XX OS SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX OS Homo sapiens.  
XX PN WO200177384-A2.  
XX PD 18-OCT-2001.  
XX PF 06-APR-2001; 2001WO-IB00713.  
XX PR 07-APR-2000; 2000DE-1019173.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX PS WPI; 2001-657177/75.

Set of oligonucleotides, useful for diagnosis and cell typing, is designed to detect single nucleotide polymorphisms and cytosine methylation status -

Claim 1: SEQ ID 268245; 29pp + Sequence Listing; German.

This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The oligonucleotides are used for diagnosis and/or prognosis of cancer and a range of diseases including immune system, gastrointestinal, respiratory, central nervous system, cardiovascular and metabolic disorders. The oligomers are also used for detecting cell type differentiation. ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and AB100010-AB182073 represent the oligomers described in the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 12 BP; 10 A; 0 C; 0 G; 2 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH68268 ..  
Align seg 1/1 to: ABH68268 from: 1 to: 12

1 LysLysAsn 3  
|||||  
2 AAAAAAAT 10

seq\_name: /SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH68290

## seq\_documentation\_block:

ID ABH68290 standard; DNA; 12 BP.

XX AC ABH68290;  
XX DT 22-FEB-2002 (first entry)  
XX DE Oligonucleotide primer SEQ ID NO 268267 for detecting SNP TSC0001028.  
XX OS SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX OS Homo sapiens.  
XX PN WO200177384-A2.  
XX PD 18-OCT-2001.  
XX PF 06-APR-2001; 2001WO-IB00713.  
XX PR 07-APR-2000; 2000DE-1019173.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX PS WPI; 2001-657177/75.

Set of oligonucleotides, useful for diagnosis and cell typing, is designed to detect single nucleotide polymorphisms and cytosine methylation status -

Claim 1: SEQ ID 268267; 29pp + Sequence Listing; German.

This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The oligonucleotides are used for diagnosis and/or prognosis of cancer and a range of diseases including immune system, gastrointestinal, respiratory, central nervous system, cardiovascular and metabolic disorders. The oligomers are also used for detecting cell type differentiation. ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and AB100010-AB182073 represent the oligomers described in the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 12 BP; 10 A; 1 C; 0 G; 1 T; 0 other;

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x ABH68290 ..

Align seg 1/1 to: ABH68290 from: 1 to: 12

1 LysLysAsn 3  
|||||



2 AAAAAAAAAAC 10







OM of: US-09-528-682-2 to: EST:\* out\_format : pfs

Date: Jun 18, 2002 6:25 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool/US09528682/runat\_18062002\_082443\_7797/app\_query.fasta\_1.689  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=1000 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09528682\_@C@N1\_1\_5309  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-2

Query length: 3

Database: EST:\*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 4056.470000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_estl:AW245664	-	90.14	9.0e+04	17	AW245664 2822994.3prime NIH_MGC
gb_estl:AW248747	-	89.48	9.7e+04	19	AW248747 2821119.3prime NIH_MGC
gb_gss:AZ357587	+	89.48	9.7e+04	19	AZ357587 1M0099C15F Mouse 10kb F
gb_gss:AZ357587	+	89.48	9.7e+04	19	AZ357587 1M0099C15F Mouse 10kb F
gb_gss:AZ358099	+	89.48	9.7e+04	19	AZ358099 1M0100N12F Mouse 10kb F
gb_gss:AZ358099	+	89.48	9.7e+04	19	AZ358099 1M0100N12F Mouse 10kb F
gb_gss:AZ363907	+	89.48	9.7e+04	19	AZ363907 1M010113R Mouse 10kb F
gb_gss:AZ359929	+	89.48	9.7e+04	19	AZ359929 1M035A407F Mouse 10kb F
gb_gss:AZ514792	+	89.48	9.7e+04	19	AZ514792 1M0361B17R Mouse 10kb F
gb_gss:AZ611602	+	89.48	9.7e+04	19	AZ611602 1M0438G03F Mouse 10kb F
gb_gss:AZ654747	+	89.48	9.7e+04	19	AZ654747 1M0529F08F Mouse 10kb F
gb_gss:AZ764506	+	89.48	9.7e+04	19	AZ764506 1M0560N04R Mouse 10kb F
gb_gss:AZ764506	+	89.48	9.7e+04	19	AZ764506 1M0560N04R Mouse 10kb F
gb_gss:AZ764517	+	89.48	9.7e+04	19	AZ764517 1M0560L07R Mouse 10kb F
gb_gss:AZ766990	+	89.48	9.7e+04	19	AZ766990 1M0564H19R Mouse 10kb F
gb_gss:AZ777642	+	89.48	9.7e+04	19	AZ777642 2M0012M20F Mouse 10kb F
gb_gss:AZ796963	+	89.48	9.7e+04	19	AZ796963 2M005X2K2R Mouse 10kb F
gb_gss:AZ933338	+	89.48	9.7e+04	19	AZ933338 2M0218104R Mouse 10kb F
gb_estl:AL587630	-	89.18	1.0e+05	20	AL587630 1M0587630 BP Chicken Bra
gb_estl:AL587727	-	89.18	1.0e+05	20	AL587727 1M0587727 BP Chicken Bra
gb_estl:AL587759	-	89.18	1.0e+05	20	AL587759 1M0587759 BP Chicken Bra
gb_gss:AZ310481	+	89.18	1.0e+05	20	AZ310481 1M0025O23F Mouse 10kb F
gb_gss:AZ316368	+	89.18	1.0e+05	20	AZ316368 1M0034L22F Mouse 10kb F
gb_gss:AZ317019	+	89.18	1.0e+05	20	AZ317019 1M0073P03R Mouse 10kb F
gb_gss:AZ31237	+	89.18	1.0e+05	20	AZ31237 1M0073P03R Mouse 10kb F
gb_gss:AZ368518	+	89.18	1.0e+05	20	AZ368518 1M0118G09R Mouse 10kb F
gb_gss:AZ375620	+	89.18	1.0e+05	20	AZ375620 1M0129A08F Mouse 10kb F
gb_gss:AZ427740	+	89.18	1.0e+05	20	AZ427740 1M0205J23R Mouse 10kb F
gb_gss:AZ486787	+	89.18	1.0e+05	20	AZ486787 1M0347G11F Mouse 10kb F
gb_gss:AZ506216	+	89.18	1.0e+05	20	AZ506216 1M038A112R Mouse 10kb F
gb_gss:AZ584232	+	89.18	1.0e+05	20	AZ584232 1M038A112R Mouse 10kb F
gb_gss:AZ633741	+	89.18	1.0e+05	20	AZ633741 1M0489G12F Mouse 10kb F
gb_gss:AZ764493	+	89.18	1.0e+05	20	AZ764493 1M0560E04R Mouse 10kb F
gb_gss:AZ764493	+	89.18	1.0e+05	20	AZ764493 1M0560E04R Mouse 10kb F
gb_gss:AZ764514	+	89.18	1.0e+05	20	AZ764514 1M0560F09R Mouse 10kb F
gb_gss:AZ773905	+	89.18	1.0e+05	20	AZ773905 2M0001C15R Mouse 10kb F
gb_gss:AZ786781	+	89.18	1.0e+05	20	AZ786781 2M0032C05R Mouse 10kb F
gb_gss:AZ789280	+	89.18	1.0e+05	20	AZ789280 2M0037A08F Mouse 10kb F
gb_gss:AZ816496	+	89.18	1.0e+05	20	AZ816496 2M0085G15F Mouse 10kb F

gb_gss:AZ832043	-	16.00	89.18	1.0e+05	20	AZ832043 2M0112111F Mouse 10k
gb_gss:AZ862308	-	16.00	89.18	1.0e+05	20	AZ862308 2M0160Q20F Mouse 10k
gb_gss:AZ943013	+	16.00	89.18	1.0e+05	20	AZ943013 2M0203C01R Mouse 10k
gb_gss:AZ949997	-	16.00	89.18	1.0e+05	20	AZ949997 2M0213D02R Mouse 10k
gb_gss:AZ969913	-	16.00	89.18	1.0e+05	20	AZ969913 2M0242C22R Mouse 10k
gb_estl:AW248782	-	16.00	88.89	1.1e+05	21	AW248782 2821017.3prime NIH_M
gb_estl:AW248782	-	16.00	88.89	1.1e+05	21	AW248782 2821017.3prime NIH_M
gb_gss:AZ308846	-	16.00	88.89	1.1e+05	21	AZ308846 1M0012H15F Mouse 10k
gb_gss:AZ308846	-	16.00	88.89	1.1e+05	21	AZ308846 1M0012H15F Mouse 10k
gb_gss:AZ309774	-	16.00	88.89	1.1e+05	21	AZ309774 1M0017F01F Mouse 10k
gb_gss:AZ316019	+	16.00	88.89	1.1e+05	21	AZ316019 1M0033C05R Mouse 10k
gb_gss:AZ3316019	+	16.00	88.89	1.1e+05	21	AZ3316019 1M0059K09R Mouse 10k
gb_gss:AZ461824	+	16.00	88.89	1.1e+05	21	AZ461824 1M0267B23R Mouse 10k
gb_gss:AZ482045	+	16.00	88.89	1.1e+05	21	AZ482045 1M0306D24R Mouse 10k
gb_gss:AZ625662	+	16.00	88.89	1.1e+05	21	AZ625662 1M0465C23F Mouse 10k
gb_gss:AZ627840	+	16.00	88.89	1.1e+05	21	AZ627840 1M0474B05F Mouse 10k
gb_gss:AZ818565	+	16.00	88.89	1.1e+05	21	AZ818565 2M0088K10R Mouse 10k
gb_gss:AZ822653	+	16.00	88.89	1.1e+05	21	AZ822653 2M0096G01F Mouse 10k
gb_gss:AZ822653	+	16.00	88.89	1.1e+05	21	AZ822653 2M0096G01F Mouse 10k
gb_gss:AZ822825	+	16.00	88.89	1.1e+05	21	AZ822825 2M0096G120F Mouse 10k
gb_gss:AZ832496	+	16.00	88.89	1.1e+05	21	AZ832496 2M0112L24R Mouse 10k
gb_gss:AZ853429	+	16.00	88.89	1.1e+05	21	AZ853429 2M0156D09R Mouse 10k
gb_gss:TA331B12Q	+	16.00	88.89	1.1e+05	21	TA331B12Q T. brucei sheared ge
gb_estl:AA888191	+	16.00	88.62	1.1e+05	22	AA888191 of82a05.sl NCI_CGAP
gb_estl:AA888191	+	16.00	88.62	1.1e+05	22	AA888191 of82a05.sl NCI_CGAP
gb_estl:AA903498	-	16.00	88.62	1.1e+05	22	AA903498 ok62912.sl NCI_CGAP
gb_estl:AA934289	-	16.00	88.62	1.1e+05	22	AA934289 SMOVL3CAN12G03 Oncho
gb_gss:AZ307559	-	16.00	88.62	1.1e+05	22	AZ307559 1M0009O20F Mouse 10k
gb_gss:AZ307559	-	16.00	88.62	1.1e+05	22	AZ307559 1M0009O20F Mouse 10k
gb_gss:AZ316361	+	16.00	88.62	1.1e+05	22	AZ316361 1M0034I16F Mouse 10k
gb_gss:AZ344236	+	16.00	88.62	1.1e+05	22	AZ344236 1M0078F09F Mouse 10k
gb_gss:AZ442146	+	16.00	88.62	1.1e+05	22	AZ442146 1M0234B16R Mouse 10k
gb_gss:AZ471736	+	16.00	88.62	1.1e+05	22	AZ471736 1M0286G112R Mouse 10k
gb_gss:AZ589449	+	16.00	88.62	1.1e+05	22	AZ589449 1M0398M24F Mouse 10k
gb_gss:AZ764528	+	16.00	88.62	1.1e+05	22	AZ764528 1M0560L14R Mouse 10k
gb_gss:AZ764528	+	16.00	88.62	1.1e+05	22	AZ764528 1M0560L14R Mouse 10k
gb_gss:AZ781736	+	16.00	88.62	1.1e+05	22	AZ781736 2M0021K14F Mouse 10k
gb_gss:AZ786402	-	16.00	88.62	1.1e+05	22	AZ786402 2M0031G19R Mouse 10k
gb_gss:AZ822208	-	16.00	88.62	1.1e+05	22	AZ822208 2M0095P08F Mouse 10k
gb_gss:AZ843648	+	16.00	88.62	1.1e+05	22	AZ843648 2M0142B07R Mouse 10k
gb_gss:AZ943347	+	16.00	88.62	1.1e+05	22	AZ943347 2M0212K17R Mouse 10k
gb_gss:TA132E12P	-	16.00	88.62	1.1e+05	22	TA132E12P T. brucei sheared ge
gb_gss:TA160D07Q	+	16.00	88.62	1.1e+05	22	TA160D07Q T. brucei sheared ge
gb_gss:TA181E11P	+	16.00	88.62	1.1e+05	22	TA181E11P T. brucei sheared ge
gb_gss:TA193A07P	+	16.00	88.62	1.1e+05	22	TA193A07P T. brucei sheared ge
gb_gss:TA214D07P	-	16.00	88.62	1.1e+05	22	TA214D07P T. brucei sheared ge
gb_gss:TA282D06P	+	16.00	88.62	1.1e+05	22	TA282D06P T. brucei sheared ge
gb_gss:TA381D07P	+	16.00	88.62	1.1e+05	22	TA381D07P T. brucei sheared ge
gb_estl:AW334124	+	16.00	88.36	1.1e+05	23	AW334124 S30G5 AGS-1 Pneumoc
gb_gss:AZ308447	+	16.00	88.36	1.1e+05	23	AZ308447 1M0011J24F Mouse 10k
gb_gss:AZ321671	+	16.00	88.36	1.1e+05	23	AZ321671 1M0042017F Mouse 10k
gb_gss:AZ336455	+	16.00	88.36	1.1e+05	23	AZ336455 1M0066G17R Mouse 10k
gb_gss:AZ339880	+	16.00	88.36	1.1e+05	23	AZ339880 1M0077A01R Mouse 10k
gb_gss:AZ371086	+	16.00	88.36	1.1e+05	23	AZ371086 1M0122D02F Mouse 10k
gb_gss:AZ380872	+	16.00	88.36	1.1e+05	23	AZ380872 1M0137H02F Mouse 10k
gb_gss:AZ380872	+	16.00	88.36	1.1e+05	23	AZ380872 1M0137H02F Mouse 10k
gb_gss:AZ382013	+	16.00	88.36	1.1e+05	23	AZ382013 1M0138M14R Mouse 10k
gb_gss:AZ382429	+	16.00	88.36	1.1e+05	23	AZ382429 1M0139E11R Mouse 10k
gb_gss:AZ423771	+	16.00	88.36	1.1e+05	23	AZ423771 1M0203G24F Mouse 10k
gb_gss:AZ435597	+	16.00	88.36	1.1e+05	23	AZ435597 1M0222P09R Mouse 10k
gb_gss:AZ579583	+	16.00	88.36	1.1e+05	23	AZ579583 1M0367N03F Mouse 10k
gb_gss:AZ621676	+	16.00	88.36	1.1e+05	23	AZ621676 1M0455J07F Mouse 10k
gb_gss:AZ627841	+	16.00	88.36	1.1e+05	23	AZ627841 1M0474E01F Mouse 10k
gb_gss:AZ771221	+	16.00	88.36	1.1e+05	23	AZ771221 1M0573A16F Mouse 10k
gb_gss:AZ782565	+	16.00	88.36	1.1e+05	23	AZ782565 2M0023J03R Mouse 10k
gb_gss:AZ817623	+	16.00	88.36	1.1e+05	23	AZ817623 2M0087N09F Mouse 10k
gb_gss:AZ818518	+	16.00	88.36	1.1e+05	23	AZ818518 2M0088A08R Mouse 10k
gb_gss:AZ862224	+	16.00	88.36	1.1e+05	23	AZ862224 2M0169N18F Mouse 10k
gb_gss:AZ979442	+	16.00	88.36	1.1e+05	23	AZ979442 2M0256F07F Mouse 10k
gb_gss:AZ984045	-	16.00	88.36	1.1e+05	23	AZ984045 2M0265J23F Mouse 10k
gb_gss:TA2998706	+	16.00	88.36	1.1e+05	23	TA2998706 2M0285J13R Mouse 10k
gb_gss:TA2998706	+	16.00	88.36	1.1e+05	23	TA2998706 2M0285J13R Mouse 10k
gb_gss:TA279B01P	+	16.00	88.36	1.1e+05	23	TA279B01P T. brucei sheared ge
gb_gss:TA328D05P	+	16.00	88.36	1.1e+05	23	TA328D05P T. brucei sheared ge



gb_gss:TA328D05P	-	16.00	88.36	1.1e+05	23	AL493117	T. brucei sheared genom	gb_gss:TA339270	16.00	87.41	1.3e+05	27	AZ339270	IM0070K10R	Mouse	10k
gb_gss:TA370G06P	+	16.00	88.36	1.1e+05	23	AL496588	T. brucei sheared genom	gb_gss:TA2427734	16.00	87.41	1.3e+05	27	AZ2427734	IM020919R	Mouse	10k
gb_gss:TA35C06P	+	16.00	88.36	1.1e+05	23	AL455778	T. brucei sheared genom	gb_gss:TA2434283	16.00	87.41	1.3e+05	27	AZ434283	IM020980R	Mouse	10k
gb_estl:AL588263	-	16.00	88.10	1.2e+05	24	AL588263	BP Chicken Bra	gb_gss:TA2445466	16.00	87.41	1.3e+05	27	AZ445466	IM0241L16F	Mouse	10k
gb_estl:AW247159	+	16.00	88.10	1.2e+05	24	AW247159	2819627.3prime NIH_MGC	gb_gss:TA2458228	16.00	87.41	1.3e+05	27	AZ58228	IM0262C12F	Mouse	10k
gb_estl:C00745	+	16.00	88.10	1.2e+05	24	C00745	HUMG00087.03prime NIH_MGC	gb_gss:TA2528108	16.00	87.41	1.3e+05	27	AZ528108	IM0476H19F	Mouse	10k
gb_gss:TA3308017	-	16.00	88.10	1.2e+05	24	AZ3308017	IM0010M05R	Mouse 10kb	16.00	87.41	1.3e+05	27	AZ812708	IM0079M03R	Mouse	10k
gb_gss:TA330762	-	16.00	88.10	1.2e+05	24	AZ330762	IM0056L12F	Mouse 10kb	16.00	87.41	1.3e+05	27	BH011509	BG02453-3prime Droso		
gb_gss:TA349008	+	16.00	88.10	1.2e+05	24	AZ349008	IM0085O13R	Mouse 10kb	16.00	87.41	1.3e+05	27	AL487522	T. brucei sheared ge		
gb_gss:TA339663	+	16.00	88.10	1.2e+05	24	AZ39663	IM0165C10R	Mouse 10kb	16.00	87.41	1.3e+05	27	AL487522	T. brucei sheared ge		
gb_gss:TA3438069	+	16.00	88.10	1.2e+05	24	AZ38069	IM0228A10F	Mouse 10kb	16.00	87.41	1.3e+05	27	AL497120	T. brucei sheared ge		
gb_gss:TA2458112	+	16.00	88.10	1.2e+05	24	AZ58112	IM0261E24R	Mouse 10kb	16.00	87.41	1.3e+05	27	AL458267	T. brucei sheared ge		
gb_gss:TA2510115	+	16.00	88.10	1.2e+05	24	AZ510115	IM0354G23F	Mouse 10kb	16.00	87.41	1.3e+05	27	AL458267	T. brucei sheared ge		
gb_gss:TA2513027	+	16.00	88.10	1.2e+05	24	AZ513027	IM0358P22R	Mouse 10kb	16.00	87.41	1.3e+05	27	AA0337442	oj04D07.s1 NCI CGAP		
gb_gss:TA2621257	+	16.00	88.10	1.2e+05	24	AZ621257	IM0454E23F	Mouse 10kb	16.00	87.19	1.3e+05	28	AA0337442	oj04D07.s1 NCI CGAP		
gb_gss:TA2627850	+	16.00	88.10	1.2e+05	24	AZ67850	IM0474N20F	Mouse 10kb	16.00	87.19	1.3e+05	28	AI039927	tx22e03.s1 NCI CGAP		
gb_gss:TA2764496	+	16.00	88.10	1.2e+05	24	AZ76496	IM0560H02R	Mouse 10kb	16.00	87.19	1.3e+05	28	AI130321	SMOVL3CAN17D09 Oncho		
gb_gss:TA2764513	+	16.00	88.10	1.2e+05	24	AZ764513	IM0360D11R	Mouse 10kb	16.00	87.19	1.3e+05	28	AI130321	SMOVL3CAN17D09 Oncho		
gb_gss:TA2764519	+	16.00	88.10	1.2e+05	24	AZ764519	IM0560P11R	Mouse 10kb	16.00	87.19	1.3e+05	28	AI1458838	tl74h07.s1 NCI CGAP		
gb_gss:TA2782507	+	16.00	88.10	1.2e+05	24	AZ782507	IM0560P12R	Mouse 10kb	16.00	87.19	1.3e+05	28	AI588880	tg28g01.s1 NCI CGAP		
gb_gss:TA2801201	+	16.00	88.10	1.2e+05	24	AZ801201	IM0059M02R	Mouse 10kb	16.00	87.19	1.3e+05	28	AI590364	ts09e12.s1 NCI CGAP		
gb_gss:TA215C030	+	16.00	88.10	1.2e+05	24	AL479238	T. brucei sheared genom	gb_estl:AI678332	16.00	87.19	1.3e+05	28	AI678332	tx22e03.s1 NCI CGAP		
gb_gss:TA356G08P	+	16.00	88.10	1.2e+05	24	AL497144	T. brucei sheared genom	gb_estl:AI688906	16.00	87.19	1.3e+05	28	AI688906	tx22e03.s1 NCI CGAP		
gb_estl:AA903841	+	16.00	87.86	1.2e+05	25	AA903841	ok64e05.s1 NCI CGAP_GC4	gb_estl:AI738806	16.00	87.19	1.3e+05	28	AI738806	ty24f05.s1 NCI CGAP		
gb_estl:AI189662	+	16.00	87.86	1.2e+05	25	AI189662	qd18h06.s1 Soares_place	gb_estl:AI791979	16.00	87.19	1.3e+05	28	AI791979	wv30h06.s1 NCI CGAP		
gb_estl:AI1318195	+	16.00	87.86	1.2e+05	25	AI1318195	ta73e09.s1 NCI CGAP_HSC	gb_estl:AA332443	16.00	87.19	1.3e+05	28	AA332443	S8E7 AGS-1 Pneumocys		
gb_estl:AI433266	+	16.00	87.86	1.2e+05	25	AI433266	tl32g02.s1 NCI CGAP_Gas	gb_estl:AA466587	16.00	87.19	1.3e+05	28	AA466587	ve20d12.r1 Soares mo		
gb_estl:AI45764	+	16.00	87.86	1.2e+05	25	AI445764	tj09g10.s1 NCI CGAP_Gas	gb_estl:AA535352	16.00	87.19	1.3e+05	28	AA535352	vb34c09.s1 Stratagene		
gb_estl:AI620546	+	16.00	87.86	1.2e+05	25	AI620546	tu95b04.s1 NCI CGAP_Gas	gb_gss:TA3357605	16.00	87.19	1.3e+05	28	TA3357605	IM0099115F	Mouse	10k
gb_estl:AI587648	+	16.00	87.86	1.2e+05	25	AL587648	AL587648 BP Chicken Bra	gb_gss:TA3357605	16.00	87.19	1.3e+05	28	TA3357605	IM0099115F	Mouse	10k
gb_estl:AI587718	+	16.00	87.86	1.2e+05	25	AL587718	AL587718 BP Chicken Bra	gb_gss:TA358038	16.00	87.19	1.3e+05	28	TA358038	IM0100F05F	Mouse	10k
gb_estl:C02553	+	16.00	87.86	1.2e+05	25	C02553	HUMG0012453 Human adult	gb_gss:TA3381403	16.00	87.19	1.3e+05	28	TA381403	IM0138A04F	Mouse	10k
gb_estl:N331150	+	16.00	87.86	1.2e+05	25	N331150	yy06g01.s1 Soares melanoc	gb_gss:TA3381403	16.00	87.19	1.3e+05	28	TA381403	IM0138A04F	Mouse	10k
gb_gss:TA3318035	+	16.00	87.86	1.2e+05	25	AI318035	IM0036L12R	Mouse 10kb	16.00	87.19	1.3e+05	28	TA367839	IM0147H22R	Mouse	10k
gb_gss:TA3339378	+	16.00	87.86	1.2e+05	25	AZ339378	IM0071C09F	Mouse 10kb	16.00	87.19	1.3e+05	28	TA367839	IM0147H22R	Mouse	10k
gb_gss:TA3381039	+	16.00	87.86	1.2e+05	25	AZ381039	IM0137N18F	Mouse 10kb	16.00	87.19	1.3e+05	28	TA366214	IM0276E18R	Mouse	10k
gb_gss:TA3386891	+	16.00	87.86	1.2e+05	25	AZ386891	IM0146A08F	Mouse 10kb	16.00	87.19	1.3e+05	28	TA381286	IM0303I24F	Mouse	10k
gb_gss:TA3393511	+	16.00	87.86	1.2e+05	25	AZ393511	IM0156B07R	Mouse 10kb	16.00	87.19	1.3e+05	28	TA35060	IM0129A11F	Mouse	10k
gb_gss:TA2442170	+	16.00	87.86	1.2e+05	25	AZ42170	IM0234H18R	Mouse 10kb	16.00	87.19	1.3e+05	28	TA369547	IM0181008R	Mouse	10k
gb_gss:TA2510128	+	16.00	87.86	1.2e+05	25	AZ510128	IM0354L23F	Mouse 10kb	16.00	87.19	1.3e+05	28	TA243259	IM0248K19F	Mouse	10k
gb_gss:TA2510144	+	16.00	87.86	1.2e+05	25	AZ510144	IM0354P24F	Mouse 10kb	16.00	87.19	1.3e+05	28	TA2973999	IM0248K19F	Mouse	10k
gb_gss:TA2764498	+	16.00	87.86	1.2e+05	25	AZ764498	IM0360K04R	Mouse 10kb	16.00	87.19	1.3e+05	28	AL473898	T. brucei sheared ge		
gb_gss:TA2764498	+	16.00	87.86	1.2e+05	25	AZ764498	IM0360K04R	Mouse 10kb	16.00	87.19	1.3e+05	28	AL473898	T. brucei sheared ge		
gb_gss:TA2774476	+	16.00	87.86	1.2e+05	25	AZ774476	IM0004A01F	Mouse 10kb	16.00	87.19	1.3e+05	28	AL453073	T. brucei sheared ge		
gb_gss:TA2812604	+	16.00	87.86	1.2e+05	25	AZ812604	IM0079F23F	Mouse 10kb	16.00	87.19	1.3e+05	28	AL491008	T. brucei sheared ge		
gb_gss:TA2855418	+	16.00	87.86	1.2e+05	25	AZ855418	IM0159M21F	Mouse 10kb	16.00	87.19	1.3e+05	28	AL455611	T. brucei sheared ge		
gb_gss:TA2867155	+	16.00	87.86	1.2e+05	25	AZ867155	IM0177E23R	Mouse 10kb	16.00	86.99	1.3e+05	29	AW059655	AHuHf bsst dnc15 aa		
gb_gss:TA154D03P	-	16.00	87.86	1.2e+05	25	AL472971	T. brucei sheared genom	gb_estl:AA698835	16.00	86.99	1.3e+05	29	AA698835	r452 non-glandular-h		
gb_estl:AL587774	-	16.00	87.63	1.2e+05	26	AL587774	AL587774 BP Chicken Bra	gb_gss:TA3335504	16.00	86.99	1.3e+05	29	TA3335504	IM0065C06F	Mouse	10k
gb_gss:TA1588210	+	16.00	87.63	1.2e+05	26	AL588210	AL588210 BP Chicken Bra	gb_gss:TA345451	16.00	86.99	1.3e+05	29	TA345451	IM0080I12F	Mouse	10k
gb_gss:TA2303994	+	16.00	87.63	1.2e+05	26	AZ303994	IM0003J22R	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ492630	IM0236B22R	Mouse	10k
gb_gss:TA2316353	+	16.00	87.63	1.2e+05	26	AZ316353	IM0034D09F	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ621725	IM0455C16F	Mouse	10k
gb_gss:TA2355083	+	16.00	87.63	1.2e+05	26	AZ355083	IM0094D14R	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ80642	IM0058G16F	Mouse	10k
gb_gss:TA2405677	+	16.00	87.63	1.2e+05	26	AZ405677	IM0174D07R	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ807660	IM0070G12R	Mouse	10k
gb_gss:TA2422541	+	16.00	87.63	1.2e+05	26	AZ422541	IM0201G14F	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ819924	IM0091A19R	Mouse	10k
gb_gss:TA24422116	+	16.00	87.63	1.2e+05	26	AZ4422116	IM0203J40R	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ825156	IM02100N08F	Mouse	10k
gb_gss:TA2485624	+	16.00	87.63	1.2e+05	26	AZ485624	IM0313H11F	Mouse 10kb	16.00	86.99	1.3e+05	29	AZ621725	IM0455C16F	Mouse	10k
gb_gss:TA2620113	+	16.00	87.63	1.2e+05	26	AZ620113	IM0452N08R	Mouse 10kb	16.00	86.99	1.3e+05	29	AL464003	T. brucei sheared ge		
gb_gss:TA2623156	+	16.00	87.63	1.2e+05	26	AZ623156	IM0460L01R	Mouse 10kb	16.00	86.99	1.3e+05	29	AL484966	T. brucei sheared ge		
gb_gss:TA2651828	+	16.00	87.63	1.2e+05	26	AZ651828	IM0522D14R	Mouse 10kb	16.00	86.79	1.4e+05	30	AL037671	Homo sapiens mRNA; E		
gb_gss:TA2764502	+	16.00	87.63	1.2e+05	26	AZ764502	IM0360L03R	Mouse 10kb	16.00	86.79	1.4e+05	30	AW248759	2820825.3prime NIH_M		
gb_gss:TA2818035	+	16.00	87.63	1.2e+05	26	AL818035	IM0087O15R	Mouse 10kb	16.00	86.79	1.4e+05	30	BJ082112	BJ082112 N1BB Mochil		
gb_gss:TA2818035	+	16.00	87.63	1.2e+05	26	AL818035	IM0087O15R	Mouse 10kb	16.00	86.79	1.4e+05	30	AZ345551	IM0080E14F	Mouse	10k
gb_gss:TA2828616	+	16.00	87.63	1.2e+05	26	AZ2828616	IM0105O07R	Mouse 10kb	16.00	86.79	1.4e+05	30	AZ345551	IM0080E14F	Mouse	10k
gb_gss:TA1238120	+	16.00	87.63	1.2e+05	26	AL463522	T. brucei sheared genom	gb_gss:TA241244	16.00	86.79	1.4e+05	30	AZ41244	IM0080G03F	Mouse	10k
gb_gss:TA1238120	+	16.00	87.63	1.2e+05	26	AL463522	T. brucei sheared genom	gb_gss:TA241244	16.00	86.79	1.4e+05	30	AZ41244	IM0080G03F	Mouse	10k
gb_gss:TA321G11P	+	16.00	87.63	1.2e+05	26	AL492371	T. brucei sheared genom	gb_gss:TA259177	16.00	86.79	1.4e+05	30	AZ59177	IM0413G01F	Mouse	10k
gb_gss:TA379G11Q	+	16.00	87.63	1.2e+05	26	AL497895	T. brucei sheared genom	gb_gss:TA2598617	16.00	86.79	1.4e+05	30	AZ598617	IM0413G01F	Mouse	10k
gb_gss:TA379G																







gb_est2:B0052720	+	16.00	86.05	1.5e+05	34	! B052720	B052720	N1BB	Mochli	nd	gb_est1:AA938016	+	16.00	85.55	1.6e+05	37	! AA938016	0078f06.s1	NCI_CGAP_
gb_est2:B0057525	-	16.00	86.05	1.5e+05	34	! B057525	B057525	N1BB	Mochli	nd	gb_est1:AA938091	+	16.00	85.55	1.6e+05	37	! AA938091	0175f07.s1	NCI_CGAP_
gb_est2:B0058673	-	16.00	86.05	1.5e+05	34	! B058673	B058673	N1BB	Mochli	nd	gb_est1:AA939234	+	16.00	85.55	1.6e+05	37	! AA939234	0975h07.s1	Soares_NF
gb_est2:B0061790	-	16.00	86.05	1.5e+05	34	! B061790	B061790	N1BB	Mochli	nd	gb_est1:AA947987	+	16.00	85.55	1.6e+05	37	! AA947987	0q58e02.s1	NCI_CGAP_
gb_est2:B0061960	-	16.00	86.05	1.5e+05	34	! B061960	B061960	N1BB	Mochli	nd	gb_est1:AA961907	+	16.00	85.55	1.6e+05	37	! AA961907	0r68d10.s1	NCI_CGAP_
gb_est2:B0062401	-	16.00	86.05	1.5e+05	34	! B062401	B062401	N1BB	Mochli	nd	gb_est1:AA972505	+	16.00	85.55	1.6e+05	37	! AA972505	0r15c01.s1	NCI_CGAP_
gb_est2:B0064029	-	16.00	86.05	1.5e+05	34	! B064029	B064029	N1BB	Mochli	nd	gb_est1:AI000163	+	16.00	85.55	1.6e+05	37	! AI000163	0s44d06.s1	NCI_CGAP_
gb_est2:B0064541	-	16.00	86.05	1.5e+05	34	! B064541	B064541	N1BB	Mochli	nd	gb_est1:AI018068	+	16.00	85.55	1.6e+05	37	! AI018068	0v64f01.s1	Soares_te
gb_est2:B0065805	-	16.00	86.05	1.5e+05	34	! B065805	B065805	N1BB	Mochli	nd	gb_est1:AI037939	+	16.00	85.55	1.6e+05	37	! AI037939	0x53b01.x1	Soares
gb_est2:B0066776	-	16.00	86.05	1.5e+05	34	! B066776	B066776	N1BB	Mochli	nd	gb_est1:AI149322	+	16.00	85.55	1.6e+05	37	! AI149322	0x76f08.x1	Soares_pl
gb_est2:B0067181	-	16.00	86.05	1.5e+05	34	! B067181	B067181	N1BB	Mochli	nd	gb_est1:AI174610	+	16.00	85.55	1.6e+05	37	! AI174610	0x47c11.s1	Gessler_W
gb_est2:D12157	+	16.00	86.05	1.5e+05	34	! D12157	HM0005258	Liver	HepG2	ce	gb_est1:AI267150	+	16.00	85.55	1.6e+05	37	! AI267150	0q36b12.x1	Stanley_F
gb_est2:H38031	+	16.00	86.05	1.5e+05	34	! H38031	YP47e10.s1	Soares	retina	ca	gb_est1:AI351618	+	16.00	85.55	1.6e+05	37	! AI351618	0r05g04.x1	Soares_to
gb_gss:AZ322675	+	16.00	86.05	1.5e+05	34	! AZ322675	1M0043D10R	Mouse	10kb	F	gb_est1:AI379566	+	16.00	85.55	1.6e+05	37	! AI379566	0t52g07.x1	Soares_Nh
gb_gss:AZ3232675	+	16.00	86.05	1.5e+05	34	! AZ3232675	1M0043D10R	Mouse	10kb	F	gb_est1:AI413139	+	16.00	85.55	1.6e+05	37	! AI413139	0r55e07.x1	Barstead
gb_gss:AZ345609	+	16.00	86.05	1.5e+05	34	! AZ345609	1M0080C22F	Mouse	10kb	F	gb_est1:AI435659	+	16.00	85.55	1.6e+05	37	! AI435659	0h53a09.x1	NCI_CGAP_
gb_gss:AZ475648	+	16.00	86.05	1.5e+05	34	! AZ475648	1M00293G05R	Mouse	10kb	F	gb_est1:AI440016	+	16.00	85.55	1.6e+05	37	! AI440016	0t198e08.x1	NCI_CGAP_
gb_gss:AZ803210	+	16.00	86.05	1.5e+05	34	! AZ803210	2M0063C02R	Mouse	10kb	F	gb_est1:AI474144	+	16.00	85.55	1.6e+05	37	! AI474144	0t005h10.x1	NCI_CGAP_
gb_gss:AZ840416	+	16.00	86.05	1.5e+05	34	! AZ840416	2M0136P19R	Mouse	10kb	F	gb_est1:AI475428	+	16.00	85.55	1.6e+05	37	! AI475428	0t183c09.x1	NCI_CGAP_
gb_gss:CN5006ML	-	16.00	86.05	1.5e+05	34	! AL065757	Drosophila	melanogaster			gb_est1:AI498217	+	16.00	85.55	1.6e+05	37	! AI498217	0t85b004.x1	NCI_CGAP_
gb_gss:TA222809P	+	16.00	86.05	1.5e+05	34	! AL480722	T. brucei	sheared	genom		gb_est1:AI500674	+	16.00	85.55	1.6e+05	37	! AI500674	0t99b01.x1	NCI_CGAP_
gb_gss:TA222809P	+	16.00	86.05	1.5e+05	34	! AL480722	T. brucei	sheared	genom		gb_est1:AI508475	+	16.00	85.55	1.6e+05	37	! AI508475	0f79f06.x1	zebrafish
gb_gss:TA68D10P	-	16.00	86.05	1.5e+05	34	! AL457502	T. brucei	sheared	genom		gb_est1:AI641545	+	16.00	85.55	1.6e+05	37	! AI641545	0f15e11.x1	zebrafish
gb_est1:AL641482	+	16.00	85.88	1.5e+05	35	! AL641482	AL641482	XGC-neurula	Sl		gb_est1:AI653824	+	16.00	85.55	1.6e+05	37	! AI653824	0y01n06.x1	NCI_CGAP_
gb_est1:AV836236	+	16.00	85.88	1.5e+05	35	! AV836236	AV836236	K. Sato	unpubl		gb_est1:AI671076	+	16.00	85.55	1.6e+05	37	! AI671076	0t21n12.x1	NCI_CGAP_
gb_est1:AV850856	+	16.00	85.88	1.5e+05	35	! AV850856	AV850856	Nori	Sato	unpubl	gb_est1:AI679350	+	16.00	85.55	1.6e+05	37	! AI679350	0t73e08.x1	NCI_CGAP_
gb_est1:AW250894	+	16.00	85.88	1.5e+05	35	! AW250894	2821328.3	Prime	NIH_MGC		gb_est1:AI718104	+	16.00	85.55	1.6e+05	37	! AI718104	0s45b08.x1	Barstead
gb_est2:B6496638	+	16.00	85.88	1.5e+05	35	! BG496638	602537978F1	NIH_MGC	59		gb_est1:AI741717	+	16.00	85.55	1.6e+05	37	! AI741717	0w22b07.x1	Soares_NS
gb_est2:BG569185	+	16.00	85.88	1.5e+05	35	! BG569185	6025688454F1	NIH_MGC	76		gb_est1:AI801336	+	16.00	85.55	1.6e+05	37	! AI801336	0t88a01.x1	NCI_CGAP_
gb_est2:B0032028	-	16.00	85.88	1.5e+05	35	! B032028	B032028	N1BB	Mochli	nd	gb_est1:AI802260	+	16.00	85.55	1.6e+05	37	! AI802260	0t36g07.x1	NCI_CGAP_
gb_est2:B0045059	+	16.00	85.88	1.5e+05	35	! B045059	B0045059	N1BB	Mochli	nd	gb_est1:AI803262	+	16.00	85.55	1.6e+05	37	! AI803262	0t38h12.x1	Soares_to
gb_est2:B0080323	+	16.00	85.88	1.5e+05	35	! B080323	B080323	N1BB	Mochli	nd	gb_est1:AI813747	+	16.00	85.55	1.6e+05	37	! AI813747	0k79a02.x1	NCI_CGAP_
gb_est2:B0080509	+	16.00	85.88	1.5e+05	35	! B080509	B080509	N1BB	Mochli	nd	gb_est1:AI918224	+	16.00	85.55	1.6e+05	37	! AI918224	0t09a03.x1	NCI_CGAP_
gb_est2:TS0295	+	16.00	85.88	1.5e+05	35	! TS0295	yb16g12.s1	Stratagene	fet		gb_est1:AI987823	+	16.00	85.55	1.6e+05	37	! AI987823	0t587823	BP_Chicken
gb_gss:AZ351309	+	16.00	85.88	1.5e+05	35	! AZ351309	1M0098K13F	Mouse	10kb	F	gb_est1:AA196222	+	16.00	85.55	1.6e+05	37	! AA196222	0z98b06.r1	Stratagen
gb_gss:AZ365817	+	16.00	85.88	1.5e+05	35	! AZ365817	1M0112G18R	Mouse	10kb	F	gb_est1:AV832716	+	16.00	85.55	1.6e+05	37	! AV832716	0V852716	K. Sato
gb_gss:AZA10442	+	16.00	85.88	1.5e+05	35	! AZA10442	1M0128I16R	Mouse	10kb	F	gb_est1:AV852302	+	16.00	85.55	1.6e+05	37	! AV852302	0V852302	Nori
gb_gss:AZA2469855	+	16.00	85.88	1.5e+05	35	! AZA246985	1M02B307R	Mouse	10kb	F	gb_est1:AV858269	+	16.00	85.55	1.6e+05	37	! AV858269	0V858269	Nori
gb_gss:AZA275373	+	16.00	85.88	1.5e+05	35	! AZA275373	1M0293E11R	Mouse	10kb	F	gb_est1:AW452427	+	16.00	85.55	1.6e+05	37	! AW452427	2820028.3	Prime
gb_gss:AZ785657	+	16.00	85.88	1.5e+05	35	! AZ785657	2M0029F17R	Mouse	10kb	F	gb_est1:AW332808	+	16.00	85.55	1.6e+05	37	! AW332808	013D3	AGS-1
gb_gss:AZ785723	+	16.00	85.88	1.5e+05	35	! AZ785723	2M0029H19R	Mouse	10kb	F	gb_est2:BG175511	+	16.00	85.55	1.6e+05	37	! BG175511	012334	778F2
gb_gss:AZ825655	+	16.00	85.88	1.5e+05	35	! AZ825655	2M0100P23R	Mouse	10kb	F	gb_est2:BG430173	+	16.00	85.55	1.6e+05	37	! BG430173	602495	159F1
gb_est1:AL587891	+	16.00	85.71	1.6e+05	36	! AL587891	AL587891	XP	Chicken	Str	gb_est2:BG569157	+	16.00	85.55	1.6e+05	37	! BG569157	6025884	18F1
gb_est1:AL639933	+	16.00	85.71	1.6e+05	36	! AL639933	AL639933	KC-neurula	Sl		gb_est2:BJ073798	+	16.00	85.55	1.6e+05	37	! BJ073798	0J073798	N1BB
gb_est1:AV63131	+	16.00	85.71	1.6e+05	36	! AA266819	mz91a02.r1	Soares	mouse		gb_est2:BJ0739838	+	16.00	85.55	1.6e+05	37	! BJ0739838	0J0739838	N1BB
gb_est1:AV673131	+	16.00	85.71	1.6e+05	36	! AV673131	AV673131	Nori	Sato	unpubl	gb_est2:BJ040638	+	16.00	85.55	1.6e+05	37	! BJ040638	0J040638	N1BB
gb_est1:AV832415	+	16.00	85.71	1.6e+05	36	! AV832415	AV832415	K. Sato	unpubl		gb_est2:BJ053161	+	16.00	85.55	1.6e+05	37	! BJ053161	0J053161	N1BB
gb_est1:AV832705	+	16.00	85.71	1.6e+05	36	! AV832705	AV832705	K. Sato	unpubl		gb_est2:BJ054011	+	16.00	85.55	1.6e+05	37	! BJ054011	0J054011	N1BB
gb_est1:AV833605	+	16.00	85.71	1.6e+05	36	! AV833605	AV833605	K. Sato	unpubl		gb_est2:BJ066419	+	16.00	85.55	1.6e+05	37	! BJ066419	0J066419	N1BB
gb_est1:AW335193	+	16.00	85.71	1.6e+05	36	! AW335193	S44C8	AGS-1	Pneumocysti		gb_est2:DL18221	+	16.00	85.55	1.6e+05	37	! DL18221	MUSGS001774	Mouse
gb_est2:BJ000288	+	16.00	85.71	1.6e+05	36	! B000288	B000288	MF01SSA	cdNA	C	gb_est2:DL18712	+	16.00	85.55	1.6e+05	37	! DL18712	MUSGS01774	Mouse
gb_est1:BJ015573	+	16.00	85.71	1.6e+05	36	! B015573	BJ015573	NF01SSA	cdNA	C	gb_est2:D270442	+	16.00	85.55	1.6e+05	37	! D270442	YH53a06	pl
gb_est1:BJ045088	+	16.00	85.71	1.6e+05	36	! B045088	BJ045088	N1BB	Mochli	nd	gb_est2:D270442	+	16.00	85.55	1.6e+05	37	! D270442	YH53a06	pl
gb_est1:BJ062811	+	16.00	85.71	1.6e+05	36	! B062811	BJ062811	N1BB	Mochli	nd	gb_est2:BF180013	+	16.00	85.55	1.6e+05	37	! BF180013	601806	360F1
gb_est1:BJ064992	+	16.00	85.71	1.6e+05	36	! B064992	BJ064992	N1BB	Mochli	nd	gb_gss:AZ453377	+	16.00	85.55	1.6e+05	37	! AZ453377	1M0092C17F	Mouse
gb_est2:BJ062688	+	16.00	85.71	1.6e+05	36	! D20688	HUMGS01664	Human	promyel	nd	gb_gss:AZ442522	+	16.00	85.55	1.6e+05	37	! AZ442522	1M0236D07R	Mouse
gb_gss:BE894682	+	16.00	85.71	1.6e+05	36	! BE894682	601435925F1	NIH_MGC	72		gb_gss:AZ442522	+	16.00	85.55	1.6e+05	37	! AZ442522	1M0236D07R	Mouse
gb_gss:AZ338246	+	16.00	85.71	1.6e+05	36	! AZ338246	1M0069C17F	Mouse	10kb	F	gb_gss:AZ595102	+	16.00	85.55	1.6e+05	37	! AZ595102	1M0407R03R	Mouse
gb_gss:AZ445440	+	16.00	85.71	1.6e+05	36	! AZ445440	1M0080F10F	Mouse	10kb	F	gb_gss:AZ595102	+	16.00	85.55	1.6e+05	37	! AZ595102	1M0407R03R	Mouse
gb_gss:AZ446375	+	16.00	85.71	1.6e+05	36	! AZ446375	1M0242G15R	Mouse	10kb	F	gb_gss:AZ824309	+	16.00	85.55	1.6e+05	37	! AZ824309		



gb\_est2:BJ060350 38 16.00 85.39 1.6e+05 38 1 BJ060350 BJ060350 NIBB Mochii nd  
gb\_est2:BJ077725 38 16.00 85.39 1.6e+05 38 1 BJ077725 BJ077725 NIBB Mochii nd  
gb\_est2:D18710 38 16.00 85.39 1.6e+05 38 1 D18710 MUSGS01772 Mouse 3'-dired  
gb\_est2:D19563 38 16.00 85.39 1.6e+05 38 1 D19563 MUSGS00966 Mouse 3'-dired  
gb\_est2:D21039 38 16.00 85.39 1.6e+05 38 1 D21039 HUNG02033 Human promyel  
gb\_est2:H57900 38 16.00 85.39 1.6e+05 38 1 H57900 YH14C029.r1 Soares fetal  
gb\_gss:AZ345452 38 16.00 85.39 1.6e+05 38 1 AZ345452 IM0080308F Mouse 10kb F  
gb\_gss:AZ474105 38 16.00 85.39 1.6e+05 38 1 AZ474105 IM0290821F Mouse 10kb F  
gb\_gss:AZ482096 38 16.00 85.39 1.6e+05 38 1 AZ482096 IM0307605F Mouse 10kb F  
gb\_gss:AZ583903 38 16.00 85.39 1.6e+05 38 1 AZ583903 IM0388803F Mouse 10kb F  
gb\_gss:AZ589720 38 16.00 85.39 1.6e+05 38 1 AZ589720 IM0398H23F Mouse 10kb F  
gb\_gss:AZ649512 38 16.00 85.39 1.6e+05 38 1 AZ649512 IM0519A07F Mouse 10kb F  
gb\_gss:AZ812461 38 16.00 85.39 1.6e+05 38 1 AZ812461 IM02709G10F Mouse 10kb F  
gb\_gss:AZ997776 38 16.00 85.39 1.6e+05 38 1 AZ997776 2M0284W19F Mouse 10kb F  
gb\_gss:TA217801P 38 16.00 85.39 1.6e+05 38 1 AL480717 T. brucei sheared genom  
gb\_gss:TA898909P 38 16.00 85.39 1.6e+05 38 1 AL479915 T. brucei sheared genom  
gb\_gss:TA26A08P 38 16.00 85.39 1.6e+05 38 1 AL483993 T. brucei sheared genom  
gb\_gss:TA731E050 38 16.00 85.39 1.6e+05 38 1 AL495923 T. brucei sheared genom  
gb\_est1:AW2471125 39 16.00 85.24 1.7e+05 39 1 AW2471125 2819715.3prime NIH\_MGC  
gb\_est1:AW248768 39 16.00 85.24 1.7e+05 39 1 AW248768 2819715.3prime NIH\_MGC  
gb\_est2:BJ032729 39 16.00 85.24 1.7e+05 39 1 BJ032729 BJ032729 NIBB Mochii nd  
gb\_est2:BJ036223 39 16.00 85.24 1.7e+05 39 1 BJ036223 BJ036223 NIBB Mochii nd  
gb\_est2:BJ045193 39 16.00 85.24 1.7e+05 39 1 BJ045193 BJ045193 NIBB Mochii nd  
gb\_est2:BJ049570 39 16.00 85.24 1.7e+05 39 1 BJ049570 BJ049570 NIBB Mochii nd  
gb\_est2:BJ057481 39 16.00 85.24 1.7e+05 39 1 BJ057481 BJ057481 NIBB Mochii nd  
gb\_est2:BJ057535 39 16.00 85.24 1.7e+05 39 1 BJ057535 BJ057535 NIBB Mochii nd  
gb\_est2:BJ058092 39 16.00 85.24 1.7e+05 39 1 BJ058092 BJ058092 NIBB Mochii nd  
gb\_est2:BJ059482 39 16.00 85.24 1.7e+05 39 1 BJ059482 BJ059482 NIBB Mochii nd  
gb\_est2:BJ064050 39 16.00 85.24 1.7e+05 39 1 BJ064050 BJ064050 NIBB Mochii nd  
gb\_est2:D18709 39 16.00 85.24 1.7e+05 39 1 D18709 MUSGS01771 Mouse 3'-dired  
gb\_est2:D20678 39 16.00 85.24 1.7e+05 39 1 D20678 HUNG01654 Human promyel  
gb\_est2:T070882 39 16.00 85.24 1.7e+05 39 1 T070882 YC49A04.s1 Stratagene liv  
gb\_est2:BE535115 39 16.00 85.24 1.7e+05 39 1 BE535115 601231231F1 NCI\_CGAP\_Ma  
gb\_est2:BF381753 39 16.00 85.24 1.7e+05 39 1 BF381753 601815982F1 NIH\_MGC\_56  
gb\_gss:AZ332532 39 16.00 85.24 1.7e+05 39 1 AZ332532 IM0061106F Mouse 10kb F  
gb\_gss:AZ346794 39 16.00 85.24 1.7e+05 39 1 AZ346794 IM0082N08F Mouse 10kb F  
gb\_gss:AZ447242 39 16.00 85.24 1.7e+05 39 1 AZ447242 IM0244015F Mouse 10kb F  
gb\_gss:AZ447242 39 16.00 85.24 1.7e+05 39 1 AZ447242 IM0244015F Mouse 10kb F  
gb\_gss:AZ510130 39 16.00 85.24 1.7e+05 39 1 AZ510130 IM0354W19F Mouse 10kb F  
gb\_gss:AZ509763 39 16.00 85.24 1.7e+05 39 1 AZ509763 IM0434W11R Mouse 10kb F  
gb\_gss:CN5004NT 39 16.00 85.24 1.7e+05 39 1 AL767924 Drosophila melanogaster  
gb\_gss:TA104D07P 39 16.00 85.24 1.7e+05 39 1 AL767924 Drosophila melanogaster  
gb\_gss:TA384C05P 39 16.00 85.24 1.7e+05 39 1 AL461558 T. brucei sheared genom  
gb\_gss:TA384C05P 39 16.00 85.24 1.7e+05 39 1 AL498661 T. brucei sheared genom  
em\_esthum:HSM001818 40 16.00 85.09 1.7e+05 40 1 AL037487 Homo sapiens mRNA: EST  
gb\_est1:AA872575 40 16.00 85.09 1.7e+05 40 1 AA872575 Oal6h08.s1 NCI\_CGAP\_GCE  
gb\_est1:AA888211 40 16.00 85.09 1.7e+05 40 1 AA888211 Cf86d11.s1 NCI\_CGAP\_Li5  
gb\_est1:AA890611 40 16.00 85.09 1.7e+05 40 1 AA890611 ak12g03.s1 Soares\_parat  
gb\_est1:AA922076 40 16.00 85.09 1.7e+05 40 1 AA922076 oh08g07.s1 NCI\_CGAP\_Cob  
gb\_est1:AA996247 40 16.00 85.09 1.7e+05 40 1 AA996247 os18c08.s1 NCI\_CGAP\_Ki6  
gb\_est1:AI003231 40 16.00 85.09 1.7e+05 40 1 AI003231 anl1g01.s1 Stratagene s  
gb\_est1:AI280648 40 16.00 85.09 1.7e+05 40 1 AI280648 qu04b12.x1 NCI\_CGAP\_Ov2  
gb\_est1:AI318679 40 16.00 85.09 1.7e+05 40 1 AI318679 ta49g09.x1 NCI\_CGAP\_Lu2  
gb\_est1:AI499237 40 16.00 85.09 1.7e+05 40 1 AI499237 to08g09.x1 NCI\_CGAP\_Ut2  
gb\_est1:AI500577 40 16.00 85.09 1.7e+05 40 1 AI500577 tn93b02.x1 NCI\_CGAP\_Ut2  
gb\_est1:AI591417 40 16.00 85.09 1.7e+05 40 1 AI591417 tn19d04.x1 NCI\_CGAP\_Brt  
gb\_est1:AI609582 40 16.00 85.09 1.7e+05 40 1 AI609582 tw28c02.x1 NCI\_CGAP\_Ov3  
gb\_est1:AI635454 40 16.00 85.09 1.7e+05 40 1 AI635454 ts65a10.x1 NCI\_CGAP\_Ki6  
gb\_est1:AI675980 40 16.00 85.09 1.7e+05 40 1 AI675980 wd08e09.x1 NCI\_CGAP\_Co3  
gb\_est1:AI689588 40 16.00 85.09 1.7e+05 40 1 AI689588 tx84g03.x1 NCI\_CGAP\_Ut4  
gb\_est1:AI697142 40 16.00 85.09 1.7e+05 40 1 AI697142 tq06b09.x1 NCI\_CGAP\_Ut3  
gb\_est1:AI718811 40 16.00 85.09 1.7e+05 40 1 AI718811 atl15g09.x1 Barstead aon  
gb\_est1:AI813869 40 16.00 85.09 1.7e+05 40 1 AI813869 wk61b07.x1 NCI\_CGAP\_Pan  
gb\_est1:AI813961 40 16.00 85.09 1.7e+05 40 1 AI813961 wk80d04.x1 NCI\_CGAP\_Pan  
gb\_est1:AI912339 40 16.00 85.09 1.7e+05 40 1 AI912339 tz25h06.x1 NCI\_CGAP\_Ut2  
gb\_est1:AI913326 40 16.00 85.09 1.7e+05 40 1 AI913326 tz76c12.x1 NCI\_CGAP\_Pan  
gb\_est1:AI971839 40 16.00 85.09 1.7e+05 40 1 AI971839 wv29a04.x1 NCI\_CGAP\_Ov1  
gb\_est1:AL449576 40 16.00 85.09 1.7e+05 40 1 AL449576 AL449576 Homo sapiens T  
gb\_est1:AL638703 40 16.00 85.09 1.7e+05 40 1 AL638703 AL638703 XGC-egg Silura  
gb\_est1:AV672870 40 16.00 85.09 1.7e+05 40 1 AV672870 AV672870 Nori Satoh unf  
gb\_est1:AV674461 40 16.00 85.09 1.7e+05 40 1 AV674461 AV674461 Nori Satoh unf  
gb\_est1:AW248755 40 16.00 85.09 1.7e+05 40 1 AW248755 2820779.3prime NIH\_MGC  
gb\_est2:BG563012 40 16.00 85.09 1.7e+05 40 1 BG563012 602581189F1 NIH\_MGC\_76

gb\_est2:BJ000805 40 16.00 85.09 1.7e+05 40 1 BJ000805 BJ000805 MF01SSA CDN  
gb\_est2:BJ060995 40 16.00 85.09 1.7e+05 40 1 BJ060995 BJ060995 NIBB Mochii nd  
gb\_est2:BJ082856 40 16.00 85.09 1.7e+05 40 1 BJ082856 BJ082856 NIBB Mochii nd  
gb\_est2:D19132 40 16.00 85.09 1.7e+05 40 1 D19132 MUSGS01348 Mouse 3'-d1  
gb\_est2:D19132 40 16.00 85.09 1.7e+05 40 1 D19132 MUSGS01348 Mouse 3'-d1  
gb\_est2:BF213125 40 16.00 85.09 1.7e+05 40 1 BF213125 601844961F1 NIH\_MGC  
gb\_est2:BF3382039 40 16.00 85.09 1.7e+05 40 1 BF3382039 601816366F1 NIH\_MGC  
gb\_gss:AZ300900 40 16.00 85.09 1.7e+05 40 1 AZ300900 EP(X)1358-3prime Dro  
gb\_gss:AZ326980 40 16.00 85.09 1.7e+05 40 1 AZ326980 IM0050A12F Mouse 10k  
gb\_gss:AZ378391 40 16.00 85.09 1.7e+05 40 1 AZ378391 IM0133E11F Mouse 10k  
gb\_gss:AZ626539 40 16.00 85.09 1.7e+05 40 1 AZ626539 IM0466A22R Mouse 10k  
gb\_gss:AZ68398 40 16.00 85.09 1.7e+05 40 1 AZ68398 BM0201A11F Mouse 10k  
gb\_gss:BH011458 40 16.00 85.09 1.7e+05 40 1 BH011458 BG02105-5prime Droso  
gb\_gss:TA217801P 40 16.00 85.09 1.7e+05 40 1 AL479714 T. brucei sheared ge  
gb\_gss:TA898909P 40 16.00 85.09 1.7e+05 40 1 AL461090 T. brucei sheared ge  
gb\_est1:AV833897 40 16.00 84.94 1.7e+05 41 1 AV833897 AV833897 K. Sato unp  
gb\_est1:AW249851 40 16.00 84.94 1.7e+05 41 1 AW249851 2821516.3prime NIH\_M  
gb\_est2:BG426793 40 16.00 84.94 1.7e+05 41 1 BG426793 602492967F1 NIH\_MGC  
gb\_est2:BJ029914 40 16.00 84.94 1.7e+05 41 1 BJ029914 BJ029914 NIBB Mochii nd  
gb\_est2:BJ058917 40 16.00 84.94 1.7e+05 41 1 BJ058917 BJ058917 NIBB Mochii nd  
gb\_est2:C01980 40 16.00 84.94 1.7e+05 41 1 C01980 HUMGS000398 Human adu  
gb\_gss:AZ339959 40 16.00 84.94 1.7e+05 41 1 AZ339959 IM0071K09R Mouse 10k  
gb\_gss:AZ356235 40 16.00 84.94 1.7e+05 41 1 AZ356235 IM0097D16F Mouse 10k  
gb\_gss:AZ464202 40 16.00 84.94 1.7e+05 41 1 AZ464202 IM0273A01R Mouse 10k  
gb\_gss:AZ494840 40 16.00 84.94 1.7e+05 41 1 AZ494840 IM0330E19F Mouse 10k  
gb\_gss:AZ598587 40 16.00 84.94 1.7e+05 41 1 AZ598587 IM0413A04R Mouse 10k  
gb\_gss:AZ663413 40 16.00 84.94 1.7e+05 41 1 AZ663413 IM0543J01F Mouse 10k  
gb\_est1:AL587588 40 16.00 84.80 1.8e+05 42 1 AL587588 AL587588 BP Chicken  
gb\_est1:AU011591 40 16.00 84.80 1.8e+05 42 1 AU011591 AU011591 Schizosacch  
gb\_est1:AU014534 40 16.00 84.80 1.8e+05 42 1 AU014534 AU014534 Schizosacch  
gb\_est1:AV833082 40 16.00 84.80 1.8e+05 42 1 AV833082 AV833082 K. Sato unp  
gb\_est1:AV833088 40 16.00 84.80 1.8e+05 42 1 AV833088 AV833088 K. Sato unp  
gb\_est1:AV847138 40 16.00 84.80 1.8e+05 42 1 AV847138 AV847138 Nori Satoh  
gb\_est1:AW249627 40 16.00 84.80 1.8e+05 42 1 AW249627 2821202.3prime NIH\_M  
gb\_est1:AW333885 40 16.00 84.80 1.8e+05 42 1 AW333885 2821202.3prime NIH\_M  
gb\_est2:BJ000051 40 16.00 84.80 1.8e+05 42 1 BJ000051 BJ000051 MF01SSA CDN  
gb\_est2:BJ000582 40 16.00 84.80 1.8e+05 42 1 BJ000582 BJ000582 MF01SSA CDN  
gb\_est2:D45795 40 16.00 84.80 1.8e+05 42 1 D45795 HUMGS03012 Human adult  
gb\_est2:H13848 40 16.00 84.80 1.8e+05 42 1 H13848 YH08e06.s1 Soares plac  
gb\_est2:H57173 40 16.00 84.80 1.8e+05 42 1 H57173 YC20f12.r1 Stratagene  
gb\_gss:AA073112 40 16.00 84.80 1.8e+05 42 1 AQ073112 EP(3)3677-3prime Dro  
gb\_gss:AZ345439 40 16.00 84.80 1.8e+05 42 1 AZ345439 IM0080F09F Mouse 10k  
gb\_gss:AZ364092 40 16.00 84.80 1.8e+05 42 1 AZ364092 IM0110H08F Mouse 10k  
gb\_gss:AZ387486 40 16.00 84.80 1.8e+05 42 1 AZ387486 IM0114E02F Mouse 10k  
gb\_gss:AZ579518 40 16.00 84.80 1.8e+05 42 1 AZ579518 IM0367G03F Mouse 10k  
gb\_gss:AZ585162 40 16.00 84.80 1.8e+05 42 1 AZ585162 IM0390H07F Mouse 10k  
gb\_gss:AZ645652 40 16.00 84.80 1.8e+05 42 1 AZ645652 IM0511P07F Mouse 10k  
gb\_gss:AZ665709 40 16.00 84.80 1.8e+05 42 1 AZ665709 IM0547K13F Mouse 10k  
gb\_gss:AZ780875 40 16.00 84.80 1.8e+05 42 1 AZ780875 2M00118H07R Mouse 10k  
gb\_gss:AZ784193 40 16.00 84.80 1.8e+05 42 1 AZ784193 2M0026E13R Mouse 10k  
gb\_gss:AZ800233 40 16.00 84.80 1.8e+05 42 1 AZ800233 2M0058F12F Mouse 10k  
gb\_gss:AZ819209 40 16.00 84.80 1.8e+05 42 1 AZ819209 2M0089N12R Mouse 10k  
gb\_gss:TA157D05P 40 16.00 84.80 1.8e+05 42 1 TA157D05P T. brucei sheared ge  
gb\_gss:TA15810P 40 16.00 84.80 1.8e+05 42 1 AL452503 T. brucei sheared ge  
gb\_gss:TA179D11P 40 16.00 84.80 1.8e+05 42 1 AL474654 T. brucei sheared ge  
gb\_gss:TA220C03Q 40 16.00 84.80 1.8e+05 42 1 AL480311 T. brucei sheared ge  
gb\_gss:TA253A05P 40 16.00 84.80 1.8e+05 42 1 AL482885 T. brucei sheared ge  
gb\_gss:TA379A06P 40 16.00 84.80 1.8e+05 42 1 AL497629 T. brucei sheared ge  
gb\_gss:TA51F06P 40 16.00 84.80 1.8e+05 42 1 AL482835 T. brucei sheared ge  
gb\_est1:AA912350 40 16.00 84.66 1.8e+05 43 1 AA912350 O198a05.s1 NCI\_CGAP  
gb\_est1:AA931447 40 16.00 84.66 1.8e+05 43 1 AA931447 om89b09.s1 NCI\_CGAP  
gb\_est1:AA968566 40 16.00 84.66 1.8e+05 43 1 AA968566 oq75e08.s1 NCI\_CGAP  
gb\_est1:AI000180 40 16.00 84.66 1.8e+05 43 1 AI000180 oq45f08.s1 NCI\_CGAP  
gb\_est1:AI053124 40 16.00 84.66 1.8e+05 43 1 AI053124 SWAKCAC13A08SK Brugl  
gb\_est1:AI249352 40 16.00 84.66 1.8e+05 43 1 AI249352 qx53e02.x1 NCI\_CGAP  
gb\_est1:AI268095 40 16.00 84.66 1.8e+05 43 1 AI268095 qv83f08.x1 NCI\_CGAP  
gb\_est1:AI361336 40 16.00 84.66 1.8e+05 43 1 AI361336 qv35g09.x1 NCI\_CGAP  
gb\_est1:AI33370 40 16.00 84.66 1.8e+05 43 1 AI33370 t139b10.x1 NCI\_CGAP  
gb\_est1:AI442438 40 16.00 84.66 1.8e+05 43 1 AI442438 sa26h06.y1 Gm-cl004  
gb\_est1:AI445136 40 16.00 84.66 1.8e+05 43 1 AI445136 t191a06.x1 NCI\_CGAP  
gb\_est1:AI445173 40 16.00 84.66 1.8e+05 43 1 AI445173 t191e12.x1 NCI\_CGAP  
gb\_est1:AI521600 40 16.00 84.66 1.8e+05 43 1 AI521600 to65d07.x1 NCI\_CGAP



gb_est1:A1524768	16.00	84.56	1.8e+05	43	I	A1524768	to36g01.x1	NCI_CGAP_Ut4	gb_gss:AZ776615	-	16.00	84.52	1.8e+05	44	AZ776615	2M010K23F	Mouse	10k	
gb_est1:A1571023	16.00	84.56	1.8e+05	43	I	A1571023	t25f03.x1	NCI_CGAP_Par	gb_gss:AZ856367	-	16.00	84.52	1.8e+05	44	AZ856367	2M0160F15R	Mouse	10k	
gb_est1:A1583078	16.00	84.66	1.8e+05	43	I	A1583078	ts03g05.x1	NCI_CGAP_Par	gb_gss:AZ950525	-	16.00	84.52	1.8e+05	44	AZ950525	2M0220O21R	Mouse	10k	
gb_est1:A1591196	16.00	84.66	1.8e+05	43	I	A1591196	tu12g05.x1	NCI_CGAP_Brd	gb_gss:AZ969520	-	16.00	84.52	1.8e+05	44	AZ969520	2M0242M15F	Mouse	10k	
gb_est1:A1597934	16.00	84.66	1.8e+05	43	I	A1597934	tu92c06.x1	NCI_CGAP_Gas	gb_est1:AL587540	-	16.00	84.39	1.9e+05	45	AL587540	AL587540	BP Chicken	10k	
gb_est1:A1610394	16.00	84.66	1.8e+05	43	I	A1610394	tp19d11.x1	NCI_CGAP_Gas	gb_est1:AL587734	-	16.00	84.39	1.9e+05	45	AL587734	AL587734	BP Chicken	10k	
gb_est1:A1669388	16.00	84.66	1.8e+05	43	I	A1669388	ty11h08.x1	NCI_CGAP_Ut3	gb_est1:AV4848952	-	16.00	84.39	1.9e+05	45	AV4848952	AV4848952	Nori Satoh	10k	
gb_est1:A1687191	16.00	84.66	1.8e+05	43	I	A1687191	tp85a11.x1	NCI_CGAP_Ut3	gb_est1:AV4851415	-	16.00	84.39	1.9e+05	45	AV4851415	AV4851415	Nori Satoh	10k	
gb_est1:A1705814	16.00	84.66	1.8e+05	43	I	A1705814	tq10h04.x1	NCI_CGAP_Ut3	gb_est1:AV4853684	-	16.00	84.39	1.9e+05	45	AV4853684	AV4853684	Nori Satoh	10k	
gb_est1:A1758314	16.00	84.66	1.8e+05	43	I	A1758314	ty06c01.x1	NCI_CGAP_Ut3	gb_est1:AV4854173	-	16.00	84.39	1.9e+05	45	AV4854173	AV4854173	Nori Satoh	10k	
gb_est1:A1799404	16.00	84.66	1.8e+05	43	I	A1799404	tu56a07.x1	NCI_CGAP_Ut3	gb_est1:AW249952	-	16.00	84.39	1.9e+05	45	AW249952	2821663	3prime NIH_M	10k	
gb_est1:A1803919	16.00	84.66	1.8e+05	43	I	A1803919	tp37609.x1	NCI_CGAP_Par	gb_est1:BJ0300882	-	16.00	84.39	1.9e+05	45	BJ0300882	602981270F1	NCI_CGAP	10k	
gb_est1:A1814142	16.00	84.66	1.8e+05	43	I	A1814142	wk63a10.x1	NCI_CGAP_Par	gb_est1:BJ000444	-	16.00	84.39	1.9e+05	45	BJ000444	BJ000444	MF01SSA CDN	10k	
gb_est1:A18587884	16.00	84.66	1.8e+05	43	I	A18587884	AL587884	BP Chicken	Br	gb_est1:BJ001124	-	16.00	84.39	1.9e+05	45	BJ001124	MF01SSA CDN	10k	
gb_est1:A1636541	16.00	84.66	1.8e+05	43	I	A1636541	AL636541	XGC-neurula	Str	gb_est1:BJ034331	-	16.00	84.39	1.9e+05	45	BJ034331	BJ034331	NTBB Mochii	10k
gb_est1:A17014529	16.00	84.66	1.8e+05	43	I	A17014529	AU014529	Schizosacchar	om	gb_est1:BJ048121	-	16.00	84.39	1.9e+05	45	BJ048121	BJ048121	NTBB Mochii	10k
gb_est1:A1825222	16.00	84.66	1.8e+05	43	I	A1825222	AV832522	K. Sato unpubl	om	gb_est1:BJ057980	-	16.00	84.39	1.9e+05	45	BJ057980	BJ057980	NTBB Mochii	10k
gb_est1:A1832677	16.00	84.66	1.8e+05	43	I	A1832677	AV832677	K. Sato unpubl	om	gb_est1:BJ060342	-	16.00	84.39	1.9e+05	45	BJ060342	BJ060342	NTBB Mochii	10k
gb_est1:A1833807	16.00	84.66	1.8e+05	43	I	A1833807	AV833807	K. Sato unpubl	om	gb_est1:T17566	-	16.00	84.39	1.9e+05	45	T17566	mps v260	The blue guys	10k
gb_est1:A1837712	16.00	84.66	1.8e+05	43	I	A1837712	Nori Satoh un	om	gb_est1:T17566	-	16.00	84.39	1.9e+05	45	T17566	mps v260	The blue guys	10k	
gb_est1:BG502943	16.00	84.66	1.8e+05	43	I	BG502943	602550609F1	NIH_MGC_61	gb_est1:BB733238	-	16.00	84.39	1.9e+05	45	BB733238	602247257F1	NIH_MGC	10k	
gb_est1:BJ000651	16.00	84.66	1.8e+05	43	I	BJ000651	B0300651	MF01SSA CDN	gb_est1:BP691166	-	16.00	84.39	1.9e+05	45	BP691166	602247257F1	NIH_MGC	10k	
gb_est1:BJ034348	16.00	84.66	1.8e+05	43	I	BJ034348	B034348	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:BJ034955	16.00	84.66	1.8e+05	43	I	BJ034955	B034955	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:BJ035115	16.00	84.66	1.8e+05	43	I	BJ035115	B035115	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:BJ038480	16.00	84.66	1.8e+05	43	I	BJ038480	B038480	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:BJ065223	16.00	84.66	1.8e+05	43	I	BJ065223	B065223	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:BJ076106	16.00	84.66	1.8e+05	43	I	BJ076106	B076106	NTBB Mochii	om	gb_gss:AZ245869	-	16.00	84.39	1.9e+05	45	AZ245869	1M0080K14R	Mouse	10k
gb_est1:DL8700	16.00	84.66	1.8e+05	43	I	DL8700	MUSG501762	Mouse	3'-direct	gb_est1:AA706139	-	16.00	84.26	1.9e+05	46	AA706139	2J80612.s1	Soares_fe	10k
gb_est1:BE969098	16.00	84.66	1.8e+05	43	I	BE969098	601650077R1	NIH_MGC_74	gb_est1:AA706139	-	16.00	84.26	1.9e+05	46	AA706139	2J80612.s1	Soares_fe	10k	
gb_gss:AZ346657	16.00	84.66	1.8e+05	43	I	AZ346657	1M0282A04F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_est1:AZ374531	16.00	84.66	1.8e+05	43	I	AZ374531	1M0127B19F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ451032	16.00	84.66	1.8e+05	43	I	AZ451032	1M0254M16F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ510132	16.00	84.66	1.8e+05	43	I	AZ510132	1M0354M21F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ586882	16.00	84.66	1.8e+05	43	I	AZ586882	1M0394F1F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ638328	16.00	84.66	1.8e+05	43	I	AZ638328	1M0498110F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ778700	16.00	84.66	1.8e+05	43	I	AZ778700	2M0014P03F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ781160	16.00	84.66	1.8e+05	43	I	AZ781160	2M0019L08F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ818423	16.00	84.66	1.8e+05	43	I	AZ818423	2M0088M21F	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ819373	16.00	84.66	1.8e+05	43	I	AZ819373	2M0089P19R	Mouse	10kb	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:AZ819378	16.00	84.66	1.8e+05	43	I	AZ819378	BG01290-Sprime	Drosophila	om	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_gss:TA47H110	16.00	84.66	1.8e+05	43	I	TA47H110	brucei	sheared genome	om	gb_est1:AA865836	-	16.00	84.26	1.9e+05	46	AA865836	oh04904.s1	NCI_CGAP	10k
gb_est1:AL449735	16.00	84.52	1.8e+05	44	I	AL449735	AL449735	Homo sapiens	44	gb_est1:AL449735	-	16.00	84.26	1.9e+05	46	AL449735	u090407.r1	Soares_ma	10k
gb_est1:AL586090	16.00	84.52	1.8e+05	44	I	AL586090	AL586090	BP Chicken	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AL586763	16.00	84.52	1.8e+05	44	I	AL586763	AL586763	Stratagene	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AL587842	16.00	84.52	1.8e+05	44	I	AL587842	AL587842	BP Chicken	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AL594306	16.00	84.52	1.8e+05	44	I	AL594306	AL594306	XGC-gastrula	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AL5954306	16.00	84.52	1.8e+05	44	I	AL5954306	AL5954306	XGC-gastrula	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AU007105	16.00	84.52	1.8e+05	44	I	AU007105	AU007105	Schizosacchar	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV833623	16.00	84.52	1.8e+05	44	I	AV833623	AV833623	K. Sato unpubl	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV833796	16.00	84.52	1.8e+05	44	I	AV833796	AV833796	K. Sato unpubl	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV834222	16.00	84.52	1.8e+05	44	I	AV834222	AV834222	K. Sato unpubl	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV847866	16.00	84.52	1.8e+05	44	I	AV847866	AV847866	Nori Satoh un	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV855178	16.00	84.52	1.8e+05	44	I	AV855178	AV855178	Nori Satoh un	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AW245950	16.00	84.52	1.8e+05	44	I	AW245950	28219506.3prime	NIH_MGC	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AW248967	16.00	84.52	1.8e+05	44	I	AW248967	2819506.3prime	NIH_MGC	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AGS-1	16.00	84.52	1.8e+05	44	I	AGS-1	AGS-1	Pneumocystis	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV32205	16.00	84.52	1.8e+05	44	I	AV32205	V32205	K. Sato unpubl	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV32610	16.00	84.52	1.8e+05	44	I	AV32610	AV32610	K. Sato unpubl	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:AV32937	16.00	84.52	1.8e+05	44	I	AV32937	AV32937	NTBB Mochii	44	gb_est1:AL423457	-	16.00	84.26	1.9e+05	46	AL423457	qf1311.s1	NCI_CGAP	10k
gb_est1:BJ050937	16.00	84.52	1.8e+05	44	I	BJ050937	BJ0509												



gb_gss:AZ588250	16.00	84.26	1.9e+05	46	! A2588250	1M0396020F	Mouse	10kb	gb_est1:AL643698	-	16.00	83.89	2.0e+05	49	! AL643698	AL643698	XGC-gastrul		
gb_gss:AZ77933	16.00	84.26	1.9e+05	46	! A27766403	2M00563023R	Mouse	10kb	gb_est1:AL649742	+	16.00	83.89	2.0e+05	49	! AL649742	AL649742	XGC-gastrul		
gb_gss:AZ77939	16.00	84.26	1.9e+05	46	! A2777939	2M0012M18R	Mouse	10kb	gb_est1:AL668512	+	16.00	83.89	2.0e+05	49	! AL668512	AL668512	directional		
gb_gss:AZ77958	16.00	84.26	1.9e+05	46	! A2778586	2M0013N16R	Mouse	10kb	gb_est1:AU053904	+	16.00	83.89	2.0e+05	49	! AU053904	AU053904	Dictyosteli		
gb_gss:AZ806877	16.00	84.26	1.9e+05	46	! A2806877	2M006914F	Mouse	10kb	gb_est1:AV832408	+	16.00	83.89	2.0e+05	49	! AV832408	AV832408	K. Sato unp		
gb_gss:AZA240H3P	16.00	84.26	1.9e+05	46	! A240H30	T. brucei	sheared genom		gb_est1:AV833094	+	16.00	83.89	2.0e+05	49	! AV833094	AV833094	K. Sato unp		
gb_est1:AL6427881	16.00	84.13	1.9e+05	47	! A6427881	AL627881	XGC-gastrula	sl	gb_est1:AV833225	+	16.00	83.89	2.0e+05	49	! AV833225	AV833225	K. Sato unp		
gb_est1:AL642997	16.00	84.13	1.9e+05	47	! A642997	AL642997	XGC-neurula	sl	gb_est1:AV860511	+	16.00	83.89	2.0e+05	49	! AV860511	AV860511	Nori Satoh		
gb_est1:AV675259	16.00	84.13	1.9e+05	47	! AV675259	AV675259	Nori Satoh	unp	gb_est1:AW432778	+	16.00	83.89	2.0e+05	49	! AW432778	sh82h02.y1	Gm-cl016		
gb_est2:BJ036309	16.00	84.13	1.9e+05	47	! BJ036309	BJ036309	NIBB	Mochii	nc	gb_est1:AA542254	+	16.00	83.89	2.0e+05	49	! AA542254	vf14c03.sl	Knowles S	
gb_est2:BJ043514	16.00	84.13	1.9e+05	47	! BJ043514	BJ043514	NIBB	Mochii	nc	gb_est1:AA515410	+	16.00	83.89	2.0e+05	49	! AA515410	nf68a09.sl	NCI_CGAP_	
gb_est2:BJ047667	16.00	84.13	1.9e+05	47	! BJ047667	BJ047667	NIBB	Mochii	nc	gb_est1:AA519644	+	16.00	83.89	2.0e+05	49	! AA519644	Tg5Strz2k2h05.sl	rgme	
gb_est2:BJ053798	16.00	84.13	1.9e+05	47	! BJ053798	BJ053798	NIBB	Mochii	nc	gb_est1:AA589132	+	16.00	83.89	2.0e+05	49	! AA589132	Y153g09.r1	Beddingto	
gb_est2:BJ078279	16.00	84.13	1.9e+05	47	! BJ078279	BJ078279	NIBB	Mochii	nc	gb_est2:BG409010	+	16.00	83.89	2.0e+05	49	! BG409010	gb85d05.y1	Moss EST	
gb_est2:C01236	16.00	84.13	1.9e+05	47	! C01236	HUMGS0007957	Human	adult	gb_est2:BG409272	+	16.00	83.89	2.0e+05	49	! BG409272	gb82f08.y1	Moss EST		
gb_est2:C21078	16.00	84.13	1.9e+05	47	! C21078	HUMGS0002592	Human	adult	gb_est2:BJ030140	+	16.00	83.89	2.0e+05	49	! BJ030140	BJ030140	NIBB	Mochii	
gb_est2:D18780	16.00	84.13	1.9e+05	47	! D18780	HUMGS000496	Mouse	3'-dired	gb_est2:BJ031525	+	16.00	83.89	2.0e+05	49	! BJ031525	BJ031525	NIBB	Mochii	
gb_est2:BE732251	16.00	84.13	1.9e+05	47	! BE732251	601568096F1	NIH_MGC_21		gb_est2:BJ040992	+	16.00	83.89	2.0e+05	49	! BJ040992	BJ040992	NIBB	Mochii	
gb_est2:BF026945	16.00	84.13	1.9e+05	47	! BF026945	601671082F1	NIH_MGC_20		gb_est2:BJ050675	+	16.00	83.89	2.0e+05	49	! BJ050675	BJ050675	NIBB	Mochii	
gb_est2:BF107886	16.00	84.13	1.9e+05	47	! BF107886	601823895R1	NIH_MGC_79		gb_est2:BJ062980	+	16.00	83.89	2.0e+05	49	! BJ062980	BJ062980	NIBB	Mochii	
gb_est2:BF143912	16.00	84.13	1.9e+05	47	! BF143912	601786439F1	NCI_CGAP_Lu		gb_est2:BJ066390	+	16.00	83.89	2.0e+05	49	! BJ066390	BJ066390	NIBB	Mochii	
gb_est2:BF573075	16.00	84.13	1.9e+05	47	! BF573075	602078712P2	NIH_MGC_52		gb_est2:BJ066390	+	16.00	83.89	2.0e+05	49	! BJ066390	BJ066390	NIBB	Mochii	
gb_gss:AQ025310	16.00	84.13	1.9e+05	47	! AQ025310	EP333133	5-prime	Drosop	gb_est2:BJ083898	+	16.00	83.89	2.0e+05	49	! BJ083898	BJ083898	NIBB	Mochii	
gb_gss:AZ345567	16.00	84.13	1.9e+05	47	! AZ345567	1M0800116F	Mouse	10kb	gb_est2:BM283646	+	16.00	83.89	2.0e+05	49	! BM283646	k150c05.y1	Ascaris s		
gb_gss:AZ433989	16.00	84.13	1.9e+05	47	! AZ433989	1M0220307F	Mouse	10kb	gb_est2:BM283347	+	16.00	83.89	2.0e+05	49	! BM283347	50071-2-9-B10.f.1	Ch		
gb_gss:AZ588937	16.00	84.13	1.9e+05	47	! AZ588937	1M0397M06R	Mouse	10kb	gb_est2:BM395447	+	16.00	83.89	2.0e+05	49	! BM395447	50072-2-9-B10.f.1	Ch		
gb_gss:AZ823735	16.00	84.13	1.9e+05	47	! AZ823735	2M0097M19R	Mouse	10kb	gb_est2:D19180	+	16.00	83.89	2.0e+05	49	! D19180	MUSGS000493	Mouse	3'-di	
gb_est1:AI004570	16.00	84.01	2.0e+05	48	! AI004570	ot54h05.sl	Soares	testi	gb_est2:D19875	+	16.00	83.89	2.0e+05	49	! D19875	HUMGS000940	Human	promy	
gb_est1:AI759157	16.00	84.01	2.0e+05	48	! AI759157	No.12	polydA/OT(-)	cold	gb_est2:HL6724	+	16.00	83.89	2.0e+05	49	! HL6724	Ym22f11.sl	Soares	infa	
gb_est1:AI587964	16.00	84.01	2.0e+05	48	! AI587964	AL587964	BP	Chicken	Br	gb_est2:BE295006	+	16.00	83.89	2.0e+05	49	! BE295006	601175396F1	NTH_MGC_	
gb_est2:BJ0667239	16.00	84.01	2.0e+05	48	! BG667239	DRABDB04	Rat	DRG	Librar	gb_est2:BE371720	+	16.00	83.89	2.0e+05	49	! BE371720	601217885F1	NCI_CGAP_	
gb_est2:BJ030281	16.00	84.01	2.0e+05	48	! BJ030281	BJ030281	NIBB	Mochii	nc	gb_est2:BF054816	+	16.00	83.89	2.0e+05	49	! BF054816	7171f04.y1	NCI_CGAP_	
gb_est2:BJ032854	16.00	84.01	2.0e+05	48	! BJ032854	BJ032854	NIBB	Mochii	nc	gb_est2:BF693288	+	16.00	83.89	2.0e+05	49	! BF693288	602249170F1	NTH_MGC_	
gb_est2:BJ040279	16.00	84.01	2.0e+05	48	! BJ040279	BJ040279	NIBB	Mochii	nc	gb_gss:AZ450094	+	16.00	83.89	2.0e+05	49	! AZ450094	1M0248J10R	Mouse	10k
gb_est2:BJ057911	16.00	84.01	2.0e+05	48	! BJ057911	BJ057911	NIBB	Mochii	nc	gb_gss:AZ514068	+	16.00	83.89	2.0e+05	49	! AZ514068	1M0360C09R	Mouse	10k
gb_est2:BJ058175	16.00	84.01	2.0e+05	48	! BJ058175	BJ058175	NIBB	Mochii	nc	gb_gss:AZ576556	+	16.00	83.89	2.0e+05	49	! AZ576556	AST-T11C0379	Genetra	
gb_est2:BJ064533	16.00	84.01	2.0e+05	48	! BJ064533	BJ064533	NIBB	Mochii	nc	gb_gss:AZ767999	+	16.00	83.89	2.0e+05	49	! AZ767999	1M0567B16R	Mouse	10k
gb_est2:BJ078372	16.00	84.01	2.0e+05	48	! BJ078372	BJ078372	NIBB	Mochii	nc	gb_gss:AZ771187	+	16.00	83.89	2.0e+05	49	! AZ771187	1M0573K07F	Mouse	10k
gb_est2:BJ083775	16.00	84.01	2.0e+05	48	! BJ083775	BJ083775	NIBB	Mochii	nc	gb_gss:AZ781187	+	16.00	83.89	2.0e+05	49	! AZ781187	2M0025G30F	Mouse	10k
gb_est2:BJ083933	16.00	84.01	2.0e+05	48	! BJ083933	BJ083933	NIBB	Mochii	nc	gb_gss:AZ783538	+	16.00	83.89	2.0e+05	49	! AZ783538	2M0025G30F	Mouse	10k
gb_est2:BJ083933	16.00	84.01	2.0e+05	48	! BJ083933	BJ083933	NIBB	Mochii	nc	gb_gss:AZ783538	+	16.00	83.89	2.0e+05	49	! AZ783538	2M0025G30F	Mouse	10k
gb_gss:AZ801189	16.00	84.01	2.0e+05	48	! AZ801189	2M0059K02R	Mouse	10kb	gb_gss:AZ8033783	+	16.00	83.89	2.0e+05	49	! AZ8033783	2M0059K02R	Mouse	10k	
gb_gss:AZ942334	16.00	84.01	2.0e+05	48	! AZ942334	2M0212B07R	Mouse	10kb	gb_gss:AZ830296	+	16.00	83.89	2.0e+05	49	! AZ830296	2M0109P24F	Mouse	10k	
gb_gss:AH626386	16.00	84.01	2.0e+05	48	! AH626386	1007113H06.1	LEL_x1	1007	gb_gss:AZ976298	+	16.00	83.89	2.0e+05	49	! AZ976298	2M0213J10R	Mouse	10k	
gb_gss:TA3048036P	16.00	84.01	2.0e+05	48	! AA490894	T. brucei	sheared genom		gb_gss:PC9303697	+	16.00	83.89	2.0e+05	49	! PC9303697	Plasmodium	chabaudi		
gb_est1:AA645079	16.00	83.89	2.0e+05	49	! AA645079	yg53f03.r1	Barstead	mouse	gb_gss:TA313H06P	+	16.00	83.89	2.0e+05	49	! TA313H06P	T. brucei	sheared ge		
gb_est1:AA757463	16.00	83.89	2.0e+05	49	! AA757463	zg41a06.sl	Soares	pinea	gb_est1:AI241767	+	16.00	83.77	2.0e+05	50	! AI241767	qu8C01.x1	NCI_CGAP_		
gb_est1:AA852759	16.00	83.89	2.0e+05	49	! AA852759	NHTSCaet15h1f1	Normal	H	gb_est1:AI307721	+	16.00	83.77	2.0e+05	50	! AI307721	tb36601.x1	NCI_CGAP_		
gb_est1:AA855774	16.00	83.89	2.0e+05	49	! AA855774	Vw74g02.r1	Stratagene	m	gb_est1:AI307851	+	16.00	83.77	2.0e+05	50	! AI307851	tb41c05.x1	NCI_CGAP_		
gb_est1:AA889285	16.00	83.89	2.0e+05	49	! AA889285	ak19d02.sl	Soares	parat	gb_est1:AA139046	+	16.00	83.77	2.0e+05	50	! AA139046	mn48c12.r1	NCI_CGAP_		
gb_est1:AA912467	16.00	83.89	2.0e+05	49	! AA912467	cm52d02.sl	NCI_CGAP	GC4	gb_est1:AI930735	+	16.00	83.77	2.0e+05	50	! AI930735	as97e06.x1	Barstead		
gb_est1:AA922891	16.00	83.89	2.0e+05	49	! AA922891	o151h05.sl	NCI_CGAP	HN3	gb_est1:AI930735	+	16.00	83.77	2.0e+05	50	! AI930735	sc46c09.y1	Gm-cl015		
gb_est1:AA998929	16.00	83.89	2.0e+05	49	! AA998929	os43f10.sl	NCI_CGAP	Br2	gb_est1:AL563789	+	16.00	83.77	2.0e+05	50	! AL563789	AL363789	ICRFP	522 a	
gb_est1:AI014639	16.00	83.89	2.0e+05	49	! AI014639	os36b09.sl	NCI_CGAP	Br2	gb_est1:AL587874	+	16.00	83.77	2.0e+05	50	! AL587874	AL587874	BP	Chicken	
gb_est1:AI056265	16.00	83.89	2.0e+05	49	! AI056265	oz02e04.sl	Soares	fetal	gb_est1:AU013302	+	16.00	83.77	2.0e+05	50	! AU013302	AU013302	Schizosacch		
gb_est1:AI084829	16.00	83.89	2.0e+05	49	! AI084829	uk62a04.y1	Soares	parat	gb_est1:AU033126	+	16.00	83.77	2.0e+05	50	! AU033126	AU033126	Sugano	Homo	
gb_est1:AI147154	16.00	83.89	2.0e+05	49	! AI147154	ok38a01.sl	Soares	NSF	gb_est1:AU102291	+	16.00	83.77	2.0e+05	50	! AU102291	AU102291	Sugano	Homo	
gb_est1:AI147508	16.00	83.89	2.0e+05	49	! AI147508	ot20f06.sl	NCI_CGAP	GC3	gb_est1:AU102293	+	16.00	83.77	2.0e+05	50	! AU102293	AU102293	Sugano	Homo	
gb_est1:AI224065	16.00	83.89	2.0e+05	49	! AI224065	qx13a08.sl	NCI_CGAP	Lym	gb_est1:AU102297	+	16.00	83.77	2.0e+05	50	! AU102297	AU102297	Sugano	Homo	
gb_est1:AI270095	16.00	83.89	2.0e+05	49	! AI270095	q63c08.sl	NCI_CGAP	Esc	gb_est1:AU102303	+	16.00	83.77	2.0e+05	50	! AU102303	AU102303	Sugano	Homo	
gb_est1:AI289496	16.00	83.89	2.0e+05	49	! AI289496	qk23a04.y1	Soares	parat	gb_est1:AU103112	+	16.00	83.77	2.0e+05	50	! AU103112	AU103112	Sugano	Homo	
gb_est1:AI316598	16.00	83.89	2.0e+05	49	! AI316598	uk62a04.y1	Schiller												



```

gb_est1:AUI04913      + 16.00  83.77  2.0e+05  50  ! AUI04913 AUI04913 Sugano Homo sa
gb_est1:AUI05362      - 16.00  83.77  2.0e+05  50  ! AUI05362 AUI05362 Sugano Homo sa
gb_est1:AUI05386      - 16.00  83.77  2.0e+05  50  ! AUI05386 AUI05386 Sugano Homo sa
gb_est1:AUI05390      - 16.00  83.77  2.0e+05  50  ! AUI05390 AUI05390 Sugano Homo sa
gb_est1:AUI05562      - 16.00  83.77  2.0e+05  50  ! AUI05562 AUI05562 Sugano Homo sa
gb_est1:AUI05572      + 16.00  83.77  2.0e+05  50  ! AUI05572 AUI05572 Sugano Homo sa
gb_est1:AUI05643      + 16.00  83.77  2.0e+05  50  ! AUI05643 AUI05643 Sugano Homo sa
gb_est1:AUI06026      - 16.00  83.77  2.0e+05  50  ! AUI06026 AUI06026 Sugano Homo sa
gb_est1:AUI06451      + 16.00  83.77  2.0e+05  50  ! AUI06451 AUI06451 Sugano Homo sa
gb_est1:AUI06514      + 16.00  83.77  2.0e+05  50  ! AUI06514 AUI06514 Sugano Homo sa
gb_est1:AUI06582      - 16.00  83.77  2.0e+05  50  ! AUI06582 AUI06582 Sugano Homo sa

```

seq\_name: gb\_est1:AW245664

```

seq_documentation_block:
LOCUS      AW245664              17 bp  mRNA  linear  EST 07-JAN-2000
DEFINITION 2822994.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822994 3',
            mRNA sequence.

```

```

ACCESSION  AW245664
VERSION    AW245664.1  GI:6588657
KEYWORDS   EST.
SOURCE     human.

```

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 17)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other\_ESTs: 2822994.5prime

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling

Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing

Project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross\_match from University of Washington Genome Center

PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley

Drosophila Genome Project. University of Washington Genome Center:

http://www.genome.washington.edu Low Quality Sequence: 0 contiguous

PHRED high quality bases following vector sequence. Very Low

Quality Sequence: Trace file contained 17 contiguous distinct peaks

following vector sequence.

Plate: LCM10 row: N column: 19.

High quality sequence stop: 19.

## FEATURES

## source

```

1. 17
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2822994"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"

```

/note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2:

ECORI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 0 a 2 c 1 g 14 t

## ORIGIN

alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AW245664/rev ..

Align seg 1/1 to reverse of: AW245664 from: 1 to: 17

1 LysLysAsn 3

13 AAAAAAAC 5

seq\_name: gb\_est1:AW248747

seq\_documentation\_block:

LOCUS AW248747

DEFINITION 2821119.3prime NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2821119 3',

mRNA sequence.

ACCESSION AW248747

VERSION AW248747.1 GI:6591740

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 19)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other\_ESTs: 2821119.5prime

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling

Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing

Project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross\_match from University of Washington Genome Center

PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley

Drosophila Genome Project. University of Washington Genome Center:

http://www.genome.washington.edu Low Quality Sequence: 7 contiguous

PHRED high quality bases following vector sequence. Very Low

Quality Sequence: Trace file contained 19 contiguous distinct peaks

following vector sequence.

Plate: LCM5 row: P column: 16

High quality sequence stop: 7.

Location/Qualifiers

1. 19

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2821119"

/clone\_lib="NIH\_MGC\_7"

/tissue\_type="small cell carcinoma"

/cell\_line="MGC3"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2:

ECORI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 1 a 0 c 0 g 18 t

## ORIGIN

alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000



## alignment\_block:

US-09-528-682-2 x AW248747/rev ..

Align seg 1/1 to reverse of: AW248747 from: 1 to: 19

1 LysLysAsn 3  
 |||||  
 19 AAAAAAAT 11

alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AZ357587 ..

Align seg 1/1 to: AZ357587 from: 1 to: 19

seq\_documentation\_block:  
 LOCUS AZ357587 19 bp DNA linear GSS 02-OCT-2000  
 DEFINITION 1M0099C15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0099C15 F, DNA sequence.

ACCESSION AZ357587  
 VERSION AZ357587.1 GI:10471287  
 KEYWORDS GSS.

## SOURCE

ORGANISM Mus musculus  
 house mouse.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 19)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0099 row: C column: 15  
 Seq primer: CGTTGTAACGACGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 19.

## FEATURES

source

1..19  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0099C15"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT  
 ORIGIN

10 a 0 c 0 g 9 t

## FEATURES

source

1..19  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0099C15"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated



with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

```
BASE COUNT      10 a      0 c      0 g      9 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x AZ357587/rev ..
Align seg 1/1 to reverse of: AZ357587 from: 1 to: 19
```

```
1 LysLysAsn 3
|||||||
18 AAAAAAAT 10
```

```
seq_name: gb_gss:AZ358099
```

```
seq_documentation_block:
LOCUS      AZ358099      19 bp      DNA      linear      GSS 02-OCT-2000
DEFINITION IM0100N12F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0100N12 F, DNA sequence.
```

```
ACCESSION      AZ358099
VERSION      AZ358099
KEYWORDS      GSS.
```

```
SOURCE      house mouse.
ORGANISM      Mus musculus.
```

```
REFERENCE
AUTHORS      Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
```

```
JOURNAL
COMMENT      Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0100 row: N column: 12
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 19.
```

```
FEATURES
source
1..19
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0100N12"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
```

polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

```
BASE COUNT      11 a      0 c      0 g      8 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x AZ358099 ..
Align seg 1/1 to: AZ358099 from: 1 to: 19
```

```
1 LysLysAsn 3
|||||||
4 AAAAAAAT 12
```

```
seq_name: gb_gss:AZ358099
```

```
seq_documentation_block:
LOCUS      AZ358099      19 bp      DNA      linear      GSS 02-OCT-2000
DEFINITION IM0100N12F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0100N12 F, DNA sequence.
```

```
ACCESSION      AZ358099
VERSION      AZ358099
KEYWORDS      GSS.
```

```
SOURCE      house mouse.
ORGANISM      Mus musculus.
```

```
REFERENCE
AUTHORS      Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
```

```
JOURNAL
COMMENT      Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0100 row: N column: 12
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 19.
```

```
FEATURES
source
1..19
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0100N12"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
```



/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 0 c 0 g 8 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ358099/rev ..

Align seg 1/1 to reverse of: AZ358099 from: 1 to: 19

1 LysLysAsn 3  
|||||  
19 AAAAAAAT 11

seq\_name: gb\_gss:AZ363907

seq\_documentation\_block:  
LOCUS AZ363907 19 bp DNA linear GSS 02-OCT-2000  
DEFINITION IM0109113R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0109113 R, DNA sequence.

ACCESSION AZ363907

VERSION AZ363907.1 GI:10477607

KEYWORDS GSS

SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 19)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0109 row: I column: 13

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

1. .19

FEATURES  
source

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0109113"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 1 a 0 c 0 g 18 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ363907/rev ..

Align seg 1/1 to reverse of: AZ363907 from: 1 to: 19

1 LysLysAsn 3  
|||||  
18 AAAAAAAT 10

seq\_name: gb\_gss:AZ509929

seq\_documentation\_block:

LOCUS AZ509929 19 bp DNA linear GSS 05-OCT-2000  
DEFINITION IM0354A07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0354A07 F, DNA sequence.

ACCESSION AZ509929

VERSION AZ509929.1 GI:10691245

KEYWORDS GSS

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 19)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu



Insert Length: 10000 Std Error: 0.00  
 Plate: 0354 row: A column: 07  
 Seq primer: CGTGTAAACGACGCGCAGT  
 Class: plasmid ends  
 High quality sequence stop: 19.  
 Location/Qualifiers

## FEATURES

source

```
1. .19
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0354A07"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /sex="Male"
```

```
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

```
BASE COUNT      17 a      0 c      0 g      2 t
ORIGIN
```

```
alignment_scores:
  Quality:      16.00      Length:      3
  Ratio:        5.333      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000
```

## alignment\_block:

US-09-528-682-2 x AZ509929 ..

Align seg 1/1 to: AZ509929 from: 1 to: 19

1 LysLysAsn 3

|||||

8 AAAAAAAT 16

seq\_name: gb\_gss:AZ514792

## seq\_documentation\_block:

```
LOCUS      AZ514792      19 bp      DNA      linear      GSS 05-OCT-2000
DEFINITION  IM0361B17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0361B17 R, DNA sequence.
```

ACCESSION AZ514792

VERSION AZ514792.1 GI:10696108

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 19)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0361 row: B column: 17  
 Seq primer: CACACAGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 19.  
 Location/Qualifiers

## FEATURES

source

```
1. .19
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0361B17"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /sex="Male"
```

```
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

```
BASE COUNT      15 a      1 c      1 g      2 t
ORIGIN
```

## alignment\_scores:

```
Quality:      16.00      Length:      3
Ratio:        5.333      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000
```

## alignment\_block:

US-09-528-682-2 x AZ514792 ..

Align seg 1/1 to: AZ514792 from: 1 to: 19

1 LysLysAsn 3

|||||

2 AAAAAAAC 10

seq\_name: gb\_gss:AZ611602

## seq\_documentation\_block:

```
LOCUS      AZ611602      19 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION  IM0438G03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0438G03 F, DNA sequence.
```

ACCESSION AZ611602

VERSION AZ611602.1 GI:11733792

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 19)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,



Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0438 row: G column: 03  
Seq primer: CGTTGTAACGACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 19.  
Location/Qualifiers  
1. 19  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U0GCLM0438G03"  
/clone\_lib="Mouse 10kb plasmid U0GCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

# FEATURES

source

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

GSS.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 19)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0529 row: F column: 08  
Seq primer: CGTTGTAACGACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 19.  
Location/Qualifiers  
1. 19  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U0GCLM0529F08"  
/clone\_lib="Mouse 10kb plasmid U0GCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

# FEATURES

source

1. 19  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U0GCLM0529F08"  
/clone\_lib="Mouse 10kb plasmid U0GCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

1 a 0 c 0 g 18 t

# alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AZ654747/rev ..  
Align seg 1/1 to reverse of: AZ654747 from: 1 to: 19

1 LysLysasn 3  
|||||  
14 AAAAAAAT 6

seq\_name: gb\_gss:AZ764506

BASE COUNT  
ORIGIN

0 a 0 c 2 g 17 t

# alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AZ611602/rev ..  
Align seg 1/1 to reverse of: AZ611602 from: 1 to: 19

1 LysLysasn 3  
|||||  
15 AAAAAAAC 7

seq\_name: gb\_gss:AZ654747

# seq\_documentation\_block:

LOCUS AZ654747 19 bp DNA linear GSS 14-DEC-2000  
DEFINITION IM0529F08F Mouse 10kb plasmid U0GCLM library Mus musculus genomic  
clone U0GCLM0529F08 F, DNA sequence.  
ACCESSION AZ654747  
VERSION AZ654747.1 GI:11791893



```

seq_documentation_block:
  19 bp      DNA      linear      GSS 16-FEB-2001
LOCUS      AZ764506
DEFINITION      IM0560N04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
                  clone UUGC1M0560N04 R, DNA sequence.
ACCESSION      AZ764506
VERSION        AZ764506.1 GI:12879539
KEYWORDS       GSS.
SOURCE         house mouse.
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 19)
AUTHORS        Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
                Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
                ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
                and Wright,D.,Weiss,R.
TITLE          Mouse whole genome scaffolding with paired end reads from 10kb
                plasmid inserts
JOURNAL        Unpublished (2000)
COMMENT        Contact: Robert B. Weiss
                University of Utah Genome Center
                Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
                84112, USA
                Tel: 801 585 5606
                Fax: 801 585 7177
                Email: ddunn@genetics.utah.edu
                Insert Length: 10000 Std Error: 0.00
                Plate: 0560 row: N column: 04
                Seq primer: CACACAGGAACACGCTATGACC
                Class: plasmid ends
                High quality sequence stop: 19.

```

# FEATURES

```

source
  1. 19
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="UUGC1M0560N04"
    /clone_lib="Mouse 10kb plasmid UUGC1M library"
    /sex="Male"
    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
    /note="Vector: PWD42nv; Purified genomic DNA from M.
    musculus C57BL/6J (male) was obtained from the Jackson
    Laboratory Mouse DNA Resource
    (http://www.jax.org/resources/documents/dnares/). The DNA
    was hydrodynamically sheared by repeated passage through a
    0.005 inch orifice at constant velocity. The sheared DNA
    was blunt end-repaired with T4 DNA polymerase and T4
    polynucleotide kinase. Adaptor oligonucleotides were
    ligated to the blunt ends in high molar excess. The
    adapted DNA was purified and size-selected for a 9.5 to
    10.5 kb range using preparative agarose gel
    electrophoresis. Vector DNA was prepared from a derivative
    of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
    inducible derivative of plasmid R1. The vector was ligated
    with adaptors complementary to the insert adaptors and
    purified. The sheared, adapted mouse DNA was annealed to
    adapted vector DNA, and transformed into
    chemically-competent E. coli XL10-Gold (Stratagene) cells
    and selected for ampicillin resistance."
BASE COUNT      11 a      0 c      0 g      8 t
ORIGIN

```

# alignment\_scores:

```

Quality: 16.00      Length: 3
Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

# alignment\_block:

```

US-09-528-682-2 x AZ764506

```

```

Align seg 1/1 to: AZ764506 from: 1 to: 19

```

```

1 LysLysAsn 3
|||||
4 AAAAAAAT 12

```

```

seq_name: gb_gss:AZ764506

```

# seq\_documentation\_block:

```

LOCUS      AZ764506
DEFINITION      19 bp      DNA      linear      GSS 16-FEB-2001
                  clone UUGC1M0560N04 R, DNA sequence.
ACCESSION      AZ764506
VERSION        AZ764506.1 GI:12879539
KEYWORDS       GSS.
SOURCE         house mouse.
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 19)
AUTHORS        Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
                Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
                ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
                and Wright,D.,Weiss,R.
TITLE          Mouse whole genome scaffolding with paired end reads from 10kb
                plasmid inserts
JOURNAL        Unpublished (2000)
COMMENT        Contact: Robert B. Weiss
                University of Utah Genome Center
                Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
                84112, USA
                Tel: 801 585 5606
                Fax: 801 585 7177
                Email: ddunn@genetics.utah.edu
                Insert Length: 10000 Std Error: 0.00
                Plate: 0560 row: N column: 04
                Seq primer: CACACAGGAACACGCTATGACC
                Class: plasmid ends
                High quality sequence stop: 19.

```

# FEATURES

```

source
  1. 19
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="UUGC1M0560N04"
    /clone_lib="Mouse 10kb plasmid UUGC1M library"
    /sex="Male"
    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
    /note="Vector: PWD42nv; Purified genomic DNA from M.
    musculus C57BL/6J (male) was obtained from the Jackson
    Laboratory Mouse DNA Resource
    (http://www.jax.org/resources/documents/dnares/). The DNA
    was hydrodynamically sheared by repeated passage through a
    0.005 inch orifice at constant velocity. The sheared DNA
    was blunt end-repaired with T4 DNA polymerase and T4
    polynucleotide kinase. Adaptor oligonucleotides were
    ligated to the blunt ends in high molar excess. The
    adapted DNA was purified and size-selected for a 9.5 to
    10.5 kb range using preparative agarose gel
    electrophoresis. Vector DNA was prepared from a derivative
    of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
    inducible derivative of plasmid R1. The vector was ligated
    with adaptors complementary to the insert adaptors and
    purified. The sheared, adapted mouse DNA was annealed to
    adapted vector DNA, and transformed into
    chemically-competent E. coli XL10-Gold (Stratagene) cells
    and selected for ampicillin resistance."
BASE COUNT      11 a      0 c      0 g      8 t
ORIGIN

```

# alignment\_scores:



Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AZ764506/rev ..

Align seg 1/1 to reverse of: AZ764506 from: 1 to: 19

1 LysLysAsn 3  
 |||||  
 19 AAAAAAAT 11

seq\_name: gb\_gss:AZ764511

## seq\_documentation\_block:

LOCUS AZ764511 19 bp DNA linear GSS 16-FEB-2001  
 DEFINITION IM0560B08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0560B08 R, DNA sequence.

## ACCESSION

AZ764511

## VERSION

AZ764511.1 GI:12879549

## KEYWORDS

GSS.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 19)

## AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.

## TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0560 row: B column: 08  
 Seq primer: CACACAGGAACACGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 19.

## FEATURES

## source

1..19  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydronically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adapted DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 13 a 0 c 0 g 6 t  
 ORIGIN

## alignment\_scores:

Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AZ764511 ..

Align seg 1/1 to: AZ764511 from: 1 to: 19

1 LysLysAsn 3  
 |||||  
 6 AAAAAAAT 14

seq\_name: gb\_gss:AZ764517

## seq\_documentation\_block:

LOCUS AZ764517 19 bp DNA linear GSS 16-FEB-2001  
 DEFINITION IM0560L07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0560L07 R, DNA sequence.

## ACCESSION

AZ764517

## VERSION

AZ764517.1 GI:12879561

## KEYWORDS

GSS.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 19)

## AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.

## TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0560 row: L column: 07  
 Seq primer: CACACAGGAACACGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 19.

## FEATURES

## source

1..19  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydronically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adapted DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adapted vector DNA, and transformed into



10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 16 a 0 c 0 g 3 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AZ764517 ..

Align seg 1/1 to: AZ764517 from: 1 to: 19

1 LysLysAsn 3  
|||||||  
9 AAAAAAAT 17

seq\_name: gb\_gss:AZ766990

seq\_documentation\_block:

LOCUS AZ766990 19 bp DNA linear GSS 16-FEB-2001  
DEFINITION LM0564H19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0564H19 R, DNA sequence.

ACCESSION AZ766990  
VERSION AZ766990.1 GI:12884624

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0564 row: H column: 19

Seq primer: CACACAGGAACGATGACC

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

1..19

source

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0564H19"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 16 a 0 c 2 g 1 t  
ORIGIN

alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AZ766990 ..

Align seg 1/1 to: AZ766990 from: 1 to: 19

1 LysLysAsn 3  
|||||||  
7 AAAAAAAT 15

seq\_name: gb\_gss:AZ777642

seq\_documentation\_block:

LOCUS AZ777642 19 bp DNA linear GSS 16-FEB-2001  
DEFINITION 2M0012M20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0012M20 F, DNA sequence.

ACCESSION AZ777642

VERSION AZ777642.1 GI:12906649

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0012 row: M column: 20

Seq primer: CGTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

1..19

source

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"



```

/clone="UUGC2M0012M20"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      0 a      0 c      10 g      9 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AZ777642/rev ..
Align seg 1/1 to reverse of: AZ777642 from: 1 to: 19

1 LysLysAsn 3
|||||||
9 AAAAAAAC 1

seq_name: gb_gss:AZ796963

seq_documentation_block:
LOCUS      AZ796963      19 bp      DNA      linear      GSS 16-FEB-2001.
DEFINITION      2M0052K24R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0052K24 R, DNA sequence.
ACCESSION      AZ796963
VERSION      AZ796963.1      GI:12945556
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000      Std Error: 0.00
Plate: 0052 row: K column: 24
Seq primer: CACACAGGAACAGCTATGACC

```

```

Class: plasmid ends
High quality sequence stop: 19.
FEATURES             source
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0052K24"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      0 a      0 c      10 g      9 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AZ796963/rev ..
Align seg 1/1 to reverse of: AZ796963 from: 1 to: 19

1 LysLysAsn 3
|||||||
18 AAAAAAAC 10

seq_name: gb_gss:AZ953338

seq_documentation_block:
LOCUS      AZ953338      19 bp      DNA      linear      GSS 27-APR-2001
DEFINITION      2M0218104R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0218104 R, DNA sequence.
ACCESSION      AZ953338
VERSION      AZ953338.1      GI:13824565
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

```



84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: dunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0218 row: 1 column: 04

Seq primer: CACACAGGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 19.

#### FEATURES

Location/Qualifiers  
1. 19

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0218104"  
/clone\_lib="Mouse 10kb plasmid UUGC2M library"  
/sex="Female"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (gll14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

#### BASE COUNT

12 a 0 c 0 g 7 t

#### ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-528-682-2 x AZ953338 ..

Align seg 1/1 to: AZ953338 from: 1 to: 19

#### 1 LysLysAsn 3

|||||

5 AAAAAAAAT 13

seq\_name: gb\_est1:AL587630

#### seq\_documentation\_block:

LOCUS AL587630 20 bp mRNA linear EST 02-MAR-2001  
DEFINITION AL587630 BP Chicken Brain Library Gallus gallus cdna clone  
ROS060A07, mRNA sequence.

#### ACCESSION

AL587630

#### VERSION

AL587630.1 GI:13192664

#### KEYWORDS

EST.

#### SOURCE

chicken.

#### ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 20)

Murray, F.

BP Chicken Brain Library

Unpublished (2001)

#### COMMENT

Contact: Frazer Murray  
Dept. Genomics and Bioinformatics  
Roslin Institute  
Roslin, Midlothian, EH25 9PS, UK  
Tel: +44 (0)131 527 4200  
Fax: +44 (0)131 440 0434  
Email: frazer.murray@bbsrc.ac.uk  
GCGGCCGCTTTT TTTT TTTT TTTT 3' Poly A RNA purchased from Clonetechn  
(\*6854-

Seq primer: M13F. Location/Qualifiers  
1. 20

#### FEATURES

Location/Qualifiers  
1. 20

/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="ROS060A07"  
/clone\_lib="BP Chicken Brain Library"

/dev\_stage="Unknown"

/tissue\_type="Brain"

/lab\_host="DH10B"

/notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; Cloned

unidirectionally. Primer: Oligo dt. 5' adaptor sequence:

5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5'

GCGGCCGCTTTT TTTT TTTT TTTT 3' Poly A RNA purchased from

Clonetechn (\*6854-1)"

BASE COUNT 2 a 2 c 1 g 15 t

ORIGIN

#### alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-528-682-2 x AL587630/rev ..

Align seg 1/1 to reverse of: AL587630 from: 1 to: 20

#### 1 LysLysAsn 3

|||||

16 AAAAAAAC 8

seq\_name: gb\_est1:AL587727

#### seq\_documentation\_block:

LOCUS AL587727 20 bp mRNA linear EST 02-MAR-2001  
DEFINITION AL587727 BP Chicken Brain Library Gallus gallus cdna clone  
ROS061D01, mRNA sequence.

#### ACCESSION

AL587727

#### VERSION

AL587727.1 GI:13192761

#### KEYWORDS

EST.

#### SOURCE

chicken.

#### ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 20)

Murray, F.

BP Chicken Brain Library

Unpublished (2001)

Contact: Frazer Murray

Dept. Genomics and Bioinformatics

Roslin Institute

Roslin, Midlothian, EH25 9PS, UK

Tel: +44 (0)131 527 4200

Fax: +44 (0)131 440 0434

Email: frazer.murray@bbsrc.ac.uk

GCGGCCGCTTTT TTTT TTTT TTTT 3' Poly A RNA purchased from Clonetechn

(\*6854-

Seq primer: M13F. Location/Qualifiers

1. 20

/organism="Gallus gallus"



```

/db_xref="taxon:9031"
/clone="ROS061D01"
/clone_lib="BP Chicken Brain Library"
/tissue_type="Brain"
/dev_stage="Unknown"
/lab_host="DH10B"
/notes="vector: pSPORT1; Site_1: NotI; Site_2: SalI; Cloned
unidirectionally. Primer: Oligo dt. 5' adaptor sequence:
5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5'
CGGCCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from
Clonetechn (*6854-1)"
BASE COUNT      1 a      0 c      1 g      18 t
ORIGIN

```

```

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-2 x AL587727/rev ..

```

```

Align seg 1/1 to reverse of: AL587727 from: 1 to: 20

```

```

1 LysLysAsn 3
|||||
9 AAAAAAAT 1

```

```

seq_name: gb_est1:AL587759

```

```

seq_documentation_block:
LOCUS AL587759 20 bp mRNA linear EST 02-MAR-2001

```

```

DEFINITION AL587759 BP Chicken Brain Library Gallus gallus cDNA clone
RS061G06, mRNA sequence.

```

```

ACCESSION AL587759
VERSION AL587759.1 GI:13192793
KEYWORDS EST.
SOURCE chicken.

```

```

ORGANISM

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 20)

```

```

AUTHORS Murray,F.

```

```

TITLE BP Chicken Brain Library

```

```

JOURNAL Unpublished (2001)

```

```

COMMENT
Contact: Frazer Murray
Dept. Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UK
Tel: +44 (0)131 527 4200
Fax: +44 (0)131 440 0434
Email: frazer.murray@bbsrc.ac.uk
CGGCCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetechn
(*6854-

```

```

Seq primer: M13F.

```

```

Location/Qualifiers

```

```

1. .20
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="ROS061G06"

```

```

/clone_lib="BP Chicken Brain Library"

```

```

/tissue_type="Brain"

```

```

/dev_stage="Unknown"

```

```

/lab_host="DH10B"

```

```

/notes="vector: pSPORT1; Site_1: NotI; Site_2: SalI; Cloned
unidirectionally. Primer: Oligo dt. 5' adaptor sequence:
5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5'
CGGCCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from
Clonetechn (*6854-1)"

```

```

BASE COUNT      2 a      0 c      0 g      18 t
ORIGIN

```

```

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-2 x AL587759/rev ..

```

```

Align seg 1/1 to reverse of: AL587759 from: 1 to: 20

```

```

1 LysLysAsn 3
|||||
9 AAAAAAAT 1

```

```

seq_name: gb_gss:AZ310481

```

```

seq_documentation_block:
LOCUS AZ310481

```

```

DEFINITION LM0025023F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0025023 F, DNA sequence.

```

```

ACCESSION AZ310481
VERSION AZ310481.1 GI:10352511
KEYWORDS GSS.
SOURCE house mouse.

```

```

ORGANISM

```

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 20)

```

```

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

```

```

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

```

```

JOURNAL Unpublished (2000)

```

```

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center

```

```

University of Utah

```

```

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

```

```

Email: ddunn@genetics.utah.edu

```

```

Insert Length: 10000 Std Error: 0.00

```

```

Plate: 0025 row: 0 column: 23

```

```

Seq primer: CGTTGTAAACGACGCGCCAGT

```

```

Class: plasmid ends

```

```

High quality sequence stop: 20.

```

```

Location/Qualifiers

```

```

source

```

```

1. .20
/organism="Mus musculus"

```

```

/strain="C57BL/6J"

```

```

/db_xref="taxon:10090"

```

```

/clone="UUGCLM0025023"

```

```

/clone_lib="Mouse 10kb plasmid UUGCLM library"

```

```

/sex="Male"

```

```

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

```

```

/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

```

```

(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated

```



with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

0 a 7 c 5 g 8 t

# alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AZ310481/rev ..

Align seg 1/1 to reverse of: AZ310481 from: 1 to: 20

1 LysLysAsn 3  
|||||||  
13 AAAAAAAC 5

seq\_name: gb\_gss:AZ316368

# seq\_documentation\_block:

LOCUS AZ316368 20 bp DNA linear GSS 29-SEP-2000  
DEFINITION 1M0034L22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0034L22 F, DNA sequence.

ACCESSION AZ316368  
VERSION AZ316368.1 GI:10364124  
KEYWORDS GSS.  
SOURCE house mouse.

# ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0034 row: L column: 22  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

# TITLE

Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

# Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0034 row: L column: 22  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

Location/Qualifiers

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0034L22"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4

# FEATURES

## SOURCE

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0034L22"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4

## FEATURES

### source

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0035P14"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 17 a 0 c 0 g 3 t  
ORIGIN

# alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-528-682-2 x AZ316368 ..

Align seg 1/1 to: AZ316368 from: 1 to: 20

1 LysLysAsn 3  
|||||||  
8 AAAAAAAT 16

seq\_name: gb\_gss:AZ317019

# seq\_documentation\_block:

LOCUS AZ317019 20 bp DNA linear GSS 29-SEP-2000  
DEFINITION 1M0035P14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0035P14 F, DNA sequence.

ACCESSION AZ317019  
VERSION AZ317019.1 GI:10365404  
KEYWORDS GSS.  
SOURCE house mouse.

# ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0035 row: P column: 14  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

# REFERENCE

## AUTHORS

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0035 row: P column: 14  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

# JOURNAL

## COMMENT

Location/Qualifiers

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0035P14"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4



/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 12 a 0 c 7 g 1 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ317019 ..  
Align seg 1/1 to: AZ317019 from: 1 to: 20

1 LysLysAsn 3  
|||||  
3 AAAAAAAT 11

seq\_name: gb\_gss:AZ341237  
seq\_documentation\_block:  
LOCUS AZ341237 20 bp DNA linear GSS 29-SEP-2000  
DEFINITION IM0073G03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0073G03 R, DNA sequence.

ACCESSION AZ341237  
VERSION AZ341237.1 GI:10417288  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112 USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0073 row: G column: 03  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers  
1. .20

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0073G03"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 19 a 0 c 0 g 1 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ341237 ..  
Align seg 1/1 to: AZ341237 from: 1 to: 20

1 LysLysAsn 3  
|||||  
12 AAAAAAAT 20

seq\_name: gb\_gss:AZ368518  
seq\_documentation\_block:  
LOCUS AZ368518 20 bp DNA linear GSS 02-OCT-2000  
DEFINITION IM0118G09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0118G09 R, DNA sequence.

ACCESSION AZ368518  
VERSION AZ368518.1 GI:10482218  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu



Insert Length: 10000 Std Error: 0.00  
 Plate: 0118 row: G column: 09  
 Seq primer: CACACAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 20.  
 Location/Qualifiers

## FEATURES

source

1. .20  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0118G09"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 16 a 1 c 0 g 3 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-528-682-2 x AZ368518 ..  
 Align seg 1/1 to: AZ368518 from: 1 to: 20  
 1 LysLysAsn 3  
 |||||  
 7 AAAAAAAC 15

seq\_name: gb\_gss:AZ375620

seq\_documentation\_block:  
 LOCUS AZ375620 20 bp DNA linear GSS 02-OCT-2000  
 DEFINITION 1M0129A08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0129A08 F, DNA sequence.

ACCESSION AZ375620  
 VERSION AZ375620.1 GI:10489320  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 20)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Isiam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.  
 and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss

University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0129 row: A column: 08  
 Seq primer: CGTGTAAACAGCGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 20.  
 Location/Qualifiers

## FEATURES

source

1. .20  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0129A08"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 19 a 0 c 0 g 1 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 16.00 Length: 3  
 Ratio: 5.333 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-528-682-2 x AZ375620 ..  
 Align seg 1/1 to: AZ375620 from: 1 to: 20  
 1 LysLysAsn 3  
 |||||  
 1 AAAAAAAC 9

seq\_name: gb\_gss:AZ427740

seq\_documentation\_block:  
 LOCUS AZ427740 20 bp DNA linear GSS 03-OCT-2000  
 DEFINITION 1M0209J23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0209J23 R, DNA sequence.

ACCESSION AZ427740  
 VERSION AZ427740.1 GI:10551753  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 20)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss



Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss  
University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0209 row: J column: 23

Seq primer: CACACGAGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 20.

#### FEATURES

Location/Qualifiers

1..20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUC1M0209J23"

/clone\_lib="Mouse 10kb plasmid UUC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

12 a

5 c

0 g

3 t

#### alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-528-682-2 x AZ427740 ..

Align seg 1/1 to: AZ427740 from: 1 to: 20

1 LysLysAsn 3

|||||

6 AAAAAAAC 14

seq\_name: gb\_gss:AZ486787

seq\_documentation\_block:

LOCUS AZ486787

DEFINITION IM0315D23F Mouse 10kb plasmid UUC1M library Mus musculus genomic

clone UUC1M0315D23 F, DNA sequence.

ACCESSION AZ486787

VERSION AZ486787.1 GI:10653904

20 bp DNA linear GSS 05-OCT-2000

#### KEYWORDS

house mouse.

#### SOURCE

Mus musculus

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 20)

#### REFERENCE

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

, and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0315 row: D column: 23

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 20.

#### FEATURES

Location/Qualifiers

1..20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUC1M0315D23"

/clone\_lib="Mouse 10kb plasmid UUC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

0 a

0 c

1 g

19 t

#### alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-528-682-2 x AZ486787/rev ..

Align seg 1/1 to reverse of: AZ486787 from: 1 to: 20

1 LysLysAsn 3

|||||

12 AAAAAAAC 4

seq\_name: gb\_gss:AZ506216



```

seq_documentation_block:
  20 bp      DNA      linear      GSS 05-OCT-2000
LOCUS      1M0347G11F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGCLM0347G11 F, DNA sequence.
ACCESSION  AZ506216
VERSION     AZ506216.1 GI:10687532
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D.,Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Insert Length: 10000 Std Error: 0.00
          Plate: 0347 row: G column: 11
          Seq primer: CGTGTAAACGACGGCCAGT
          Class: plasmid ends
          High quality sequence stop: 20.
FEATURES   Location/Qualifiers
            1..20
             /organism="Mus musculus"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="UUGCLM0347G11"
             /clone_lib="Mouse 10kb plasmid UUGCLM library"
             /sex="Male"
             /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
             /note="Vector: PWD42nv; Purified genomic DNA from M.
             Laboratory Mouse DNA Resource
             (http://www.jax.org/resources/documents/dnares/). The DNA
             was hydronamically sheared by repeated passage through a
             0.005 inch orifice at constant velocity. The sheared DNA
             was blunt end-repaired with T4 DNA polymerase and T4
             polynucleotide kinase. Adaptor oligonucleotides were
             ligated to the blunt ends in high molar excess. The
             adaptored DNA was purified and size-selected for a 9.5 to
             10.5 kb range using preparative agarose gel
             electrophoresis. Vector DNA was prepared from a derivative
             of pWD42 (gil4732114|gb|AF129072.1), a copy-number
             inducible derivative of plasmid R1. The vector was ligated
             with adaptors complementary to the insert adaptors and
             purified. The sheared, adaptored mouse DNA was annealed to
             adaptored vector DNA, and transformed into
             chemically-competent E. coli XL10-Gold (Stratagene) cells
             and selected for ampicillin resistance."
BASE COUNT      13 a      0 c      0 g      7 t
ORIGIN
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-528-682-2 x AZ506216

```

Align seg 1/1 to: AZ506216 from: 1 to: 20

```

1 LysLysAsn 3
|||||
6 AAAAAAAT 14

```

seq\_name: gb\_gss:AZ584232

```

seq_documentation_block:
  20 bp      DNA      linear      GSS 13-DEC-2000
LOCUS      1M0388A12R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGCLM0388A12 R, DNA sequence.
ACCESSION  AZ584232
VERSION     AZ584232.1 GI:11704906
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D.,Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Insert Length: 10000 Std Error: 0.00
          Plate: 0388 row: A column: 12
          Seq primer: CACACAGAAACAGCTATGACC
          Class: plasmid ends
          High quality sequence stop: 20.
FEATURES   Location/Qualifiers
            1..20
             /organism="Mus musculus"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="UUGCLM0388A12"
             /clone_lib="Mouse 10kb plasmid UUGCLM library"
             /sex="Male"
             /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
             /note="Vector: PWD42nv; Purified genomic DNA from M.
             Laboratory Mouse DNA Resource
             (http://www.jax.org/resources/documents/dnares/). The DNA
             was hydronamically sheared by repeated passage through a
             0.005 inch orifice at constant velocity. The sheared DNA
             was blunt end-repaired with T4 DNA polymerase and T4
             polynucleotide kinase. Adaptor oligonucleotides were
             ligated to the blunt ends in high molar excess. The
             adaptored DNA was purified and size-selected for a 9.5 to
             10.5 kb range using preparative agarose gel
             electrophoresis. Vector DNA was prepared from a derivative
             of pWD42 (gil4732114|gb|AF129072.1), a copy-number
             inducible derivative of plasmid R1. The vector was ligated
             with adaptors complementary to the insert adaptors and
             purified. The sheared, adaptored mouse DNA was annealed to
             adaptored vector DNA, and transformed into
             chemically-competent E. coli XL10-Gold (Stratagene) cells
             and selected for ampicillin resistance."
BASE COUNT      8 a      12 c      0 g      0 t
ORIGIN
alignment_scores:

```



Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AZ584232 ..

Align seg 1/1 to: AZ584232 from: 1 to: 20

1 LysLysasn 3  
|||||  
1 AAAAAAAC 9

seq\_name: gb\_gss:AZ633741

seq\_documentation\_block:

LOCUS AZ633741 20 bp DNA linear GSS 13-DEC-2000  
DEFINITION IM0489G12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0489G12 F, DNA sequence.

ACCESSION AZ633741

VERSION AZ633741.1 GI:11755931

KEYWORDS GSS

SOURCE house mouse.

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

CONTACT: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0489 row: G column: 12

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 20.

FEATURES

source

1..20

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0489G12"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (g14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

BASE COUNT  
ORIGIN

18 a 2 c 0 g 0 t

chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AZ633741 ..

Align seg 1/1 to: AZ633741 from: 1 to: 20

1 LysLysasn 3  
|||||  
8 AAAAAAAC 16

seq\_name: gb\_gss:AZ764493

seq\_documentation\_block:

LOCUS AZ764493

DEFINITION 1M0560E04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0560E04 R, DNA sequence.

ACCESSION AZ764493

VERSION AZ764493.1 GI:12879513

KEYWORDS GSS

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

CONTACT: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0560 row: E column: 04

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 20.

FEATURES

source

1..20

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0560E04"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to



10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 12 a 0 c 0 g 8 t

ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ764493 ..

Align seg 1/1 to: AZ764493 from: 1 to: 20

1 LysLysAsn 3  
|||||  
5 AAAAAAAT 13

seq\_name: gb\_gss:AZ764493

seq\_documentation\_block:

LOCUS AZ764493 20 bp DNA linear GSS 16-FEB-2001  
DEFINITION 1M0560E04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0560E04 R, DNA sequence.

ACCESSION AZ764493

VERSION AZ764493.1 GI:12879513

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0560 row: E column: 04

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold. T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

FEATURES

source

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 12 a 0 c 0 g 8 t

ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ764493/rev ..

Align seg 1/1 to reverse of: AZ764493 from: 1 to: 20

1 LysLysAsn 3  
|||||  
20 AAAAAAAT 12

seq\_name: gb\_gss:AZ764514

seq\_documentation\_block:

LOCUS AZ764514 20 bp DNA linear GSS 16-FEB-2001  
DEFINITION 1M0560F09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0560F09 R, DNA sequence.

ACCESSION AZ764514

VERSION AZ764514.1 GI:12879555

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0560 row: F column: 09

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

1. .20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"



```

/clone-lib="UUGC1M0560F09"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      13 a      0 c      0 g      7 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x AZ764514 ..

  Align seg 1/1 to: AZ764514 from: 1 to: 20

1 LysLysAsn 3
|||||||
6 AAAAAAAAAAT 14

seq_name: gb_gss:AZ773905

seq_documentation_block:
LOCUS      AZ773905      20 bp      DNA      linear      GSS 16-FEB-2001
DEFINITION      2M0001C15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0001C15 R, DNA sequence.
ACCESSION      AZ773905
VERSION      AZ773905.1 GI:12898761
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0001 row: C column: 15
Seq primer: CACACAGGAACAGCTATGACC

```

```

Class: plasmid ends
High quality sequence stop: 20.
FEATURES
  source
    1..20
      Location/Qualifiers
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUGC2M0001C15"
        /clone_lib="Mouse 10kb plasmid UUGC1M library"
        /sex="Male"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /notes="vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
        (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      7 a      0 c      0 g      13 t
ORIGIN

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x AZ773905/rev ..

  Align seg 1/1 to reverse of: AZ773905 from: 1 to: 20

1 LysLysAsn 3
|||||||
15 AAAAAAAAAAT 7

seq_name: gb_gss:AZ786781

seq_documentation_block:
LOCUS      AZ786781      20 bp      DNA      linear      GSS-16-FEB-2001
DEFINITION      2M0032C03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0032C05 R, DNA sequence.
ACCESSION      AZ786781
VERSION      AZ786781.1 GI:12924886
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

```



```

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0032 row: C column: 05
Seq primer: CACACAGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0032C05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
3 a 3 c 0 g 14 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AZ786781/rev ..
Align seg 1/1 to reverse of: AZ786781 from: 1 to: 20
1 LysLysAsn 3
|||||
18 AAAAAAAT 10

seq_name: gb_gss:AZ789280
seq_documentation_block:
LOCUS AZ789280 20 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0037A08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0037A08 F, DNA sequence.
ACCESSION AZ789280
VERSION AZ789280.1 GI:12930134
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0032 row: C column: 05
Seq primer: CACACAGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0032C05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
3 a 3 c 0 g 14 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AZ786781/rev ..
Align seg 1/1 to reverse of: AZ786781 from: 1 to: 20
1 LysLysAsn 3
|||||
18 AAAAAAAT 10

seq_name: gb_gss:AZ789280
seq_documentation_block:
LOCUS AZ789280 20 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0037A08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0037A08 F, DNA sequence.
ACCESSION AZ789280
VERSION AZ789280.1 GI:12930134
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE
JOURNAL
COMMENT
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0037 row: A column: 08
Seq primer: CCGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0037A08"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
13 a 4 c 0 g 3 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x AZ789280 ..
Align seg 1/1 to: AZ789280 from: 1 to: 20
1 LysLysAsn 3
|||||
9 AAAAAAAT 17

seq_name: gb_gss:AZ816496
seq_documentation_block:
LOCUS AZ816496 20 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0085G16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0085G16 F, DNA sequence.
ACCESSION AZ816496
VERSION AZ816496.1 GI:12986404
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

```



REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.  
TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0085 row: G column: 16  
Seq primer: CGTTGTAACGACGCGCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

FEATURES  
source

1. .20  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0085G16"  
/clone\_lib="Mouse 10kb plasmid UUGCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adaptor vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 0 a 3 c 3 g 14 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ816496/rev ..  
Align seg 1/1 to reverse of: AZ816496 from: 1 to: 20

1 LysLysAsn 3  
|||||||  
20 AAGAAAAAC 12  
seq\_name: gb\_gss:AZ832043

seq\_documentation\_block:  
LOCUS AZ832043 20 bp DNA linear GSS 20-FEB-2001

DEFINITION  
2M0112111F Mouse 10kb plasmid UUGCLM library Mus musculus genomic  
clone UUGC2M0112111 F, DNA sequence.  
ACCESSION  
A2832043  
VERSION  
A2832043.1 GI:13001951  
KEYWORDS  
GSS.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.  
TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0112 row: I column: 11  
Seq primer: CGTTGTAACGACGCGCAGT  
Class: plasmid ends  
High quality sequence stop: 20.

FEATURES  
source

1. .20  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0112111"  
/clone\_lib="Mouse 10kb plasmid UUGCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adaptor vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 3 a 5 c 1 g 11 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x AZ832043/rev ..  
Align seg 1/1 to reverse of: AZ832043 from: 1 to: 20  
1 LysLysAsn 3



```

|||||
12 AAGAAAAAT 4

seq_name: gb_gss:A2862308

seq_documentation_block:
LOCUS      A2862308                20 bp    DNA        linear    GSS 21-FEB-2001
DEFINITION 2M0169020F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0169020 F, DNA sequence.
ACCESSION  A2862308
VERSION    A2862308.1 GI:13059481
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0169 row: 0 column: 20
            Seq primer: CGTTGTAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES   Location/Qualifiers
            source          1..20
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="UUGC2M0169020"
                        /clone_lib="Mouse 10kb plasmid UUGC1M library"
                        /sex="Male"
                        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                        /note="vector: PWD42nv; Purified genomic DNA from M.
                        Musculus C57BL/6J (male) was obtained from the Jackson
                        Laboratory Mouse DNA Resource
                        (http://www.jax.org/resources/documents/dnares/). The DNA
                        was hydrodynamically sheared by repeated passage through a
                        0.005 inch orifice at constant velocity. The sheared DNA
                        was blunt end-repaired with T4 DNA polymerase and T4
                        polynucleotide kinase. Adaptor oligonucleotides were
                        ligated to the blunt ends in high molar excess. The
                        adaptor DNA was purified and size-selected for a 9.5 to
                        10.5 kb range using preparative agarose gel
                        electrophoresis. Vector DNA was prepared from a derivative
                        of pWD42 (gii4732114|gb|AF129072.1), a copy-number
                        inducible derivative of plasmid R1. The vector was ligated
                        with adaptors complementary to the insert adaptors and
                        purified. The sheared, adaptor mouse DNA was annealed to
                        adaptor vector DNA, and transformed into
                        chemically-competent E. coli XL10-Gold (Stratagene) cells
                        and selected for ampicillin resistance."
BASE COUNT      2 a      1 c      7 g      10 t
ORIGIN

alignment_scores:
Quality:      16.00      Length:      3
Ratio:        5.333      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x A2862308/rev
Align seg 1/1 to reverse of: A2862308 from: 1 to: 20

1 LysLysAsn 3
|||||
11 AAAAAAAGAT 3

seq_name: gb_gss:A2943013

seq_documentation_block:
LOCUS      A2943013                20 bp    DNA        linear    GSS 26-APR-2001
DEFINITION 2M0203C01R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0203C01 R, DNA sequence.
ACCESSION  A2943013
VERSION    A2943013.1 GI:13806752
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0203 row: C column: 01
            Seq primer: CACACAGGAACACGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES   Location/Qualifiers
            source          1..20
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="UUGC2M0203C01"
                        /clone_lib="Mouse 10kb plasmid UUGC2M library"
                        /sex="Female"
                        /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                        /note="Vector: PWD42nv; Purified genomic DNA from M.
                        Musculus C57BL/6J (female) was obtained from the Jackson
                        Laboratory Mouse DNA Resource
                        (http://www.jax.org/resources/documents/dnares/). The DNA
                        was hydrodynamically sheared by repeated passage through a
                        0.005 inch orifice at constant velocity. The sheared DNA
                        was blunt end-repaired with T4 DNA polymerase and T4
                        polynucleotide kinase. Adaptor oligonucleotides were
                        ligated to the blunt ends in high molar excess. The
                        adaptor DNA was purified and size-selected for a 9.5 to
                        10.5 kb range using preparative agarose gel
                        electrophoresis. Vector DNA was prepared from a derivative
                        of pWD42 (gii4732114|gb|AF129072.1), a copy-number
                        inducible derivative of plasmid R1. The vector was ligated
                        with adaptors complementary to the insert adaptors and
                        purified. The sheared, adaptor mouse DNA was annealed to
                        adaptor vector DNA, and transformed into
                        chemically-competent E. coli XL10-Gold (Stratagene) cells
                        and selected for ampicillin resistance."
BASE COUNT      11 a      9 c      0 g      0 t
ORIGIN

```



## ORIGIN

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AZ943013 ..

Align seg 1/1 to: AZ943013 from: 1 to: 20

1 LysLysAsn 3  
|||||||  
4 AAAAAAAAAAC 12

seq\_name: gb\_gss:AZ949997

## seq\_documentation\_block:

LOCUS AZ949997 20 bp DNA linear GSS 27-APR-2001  
DEFINITION 2M0213D24R Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
clone UUGC2M0213D24 R, DNA sequence.

ACCESSION AZ949997  
VERSION AZ949997.1 GI:13821224  
KEYWORDS GSS.

## SOURCE

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

## TITLE

plasmid inserts

## JOURNAL

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0213 row: D column: 24

Seq primer: CACACAGGAACAGCATGACC

Class: plasmid ends

High quality sequence stop: 20.

## FEATURES

source

1..20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0213D24"

/clone\_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (female) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi14732114|gb|AF129072.1), a copy-number

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x AZ949997 ..

Align seg 1/1 to: AZ949997 from: 1 to: 20

1 LysLysAsn 3  
|||||||  
1 AAAAAAAT 9

seq\_name: gb\_gss:AZ969913

## seq\_documentation\_block:

LOCUS AZ969913 20 bp DNA linear GSS 27-APR-2001  
DEFINITION 2M0242C22R Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
clone UUGC2M0242C22 R, DNA sequence.

ACCESSION AZ969913

VERSION AZ969913.1 GI:13841140

KEYWORDS GSS.

SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

## TITLE

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0242 row: C column: 22

Seq primer: CACACAGGAACAGCATGACC

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

1..20

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0242C22"

/clone\_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (female) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA



was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (gii4732114[gb|AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 3 a 2 c 3 g 12 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AZ969913/rev ..

Align seg 1/1 to reverse of: AZ969913 from: 1 to: 20

1 LysLysAsn 3  
|||||  
9 AAGAAAAAT 1

seq\_name: gb\_est1:AW248782

seq\_documentation\_block:

LOCUS AW248782 21 bp mRNA linear EST 07-JAN-2000  
DEFINITION 2821017.3prime NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2821017 3',  
mRNA sequence.

ACCESSION AW248782

VERSION AW248782.1 GI:6591775

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 21)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other\_ESTs: 2821017.5prime

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.biol.llnl.gov/bbrp/image/html/BaseCalling/Quality](http://www.biol.llnl.gov/bbrp/image/html/BaseCalling/Quality)

Scores: PHRED from University of Washington Genome Center. Vector Trimming: cross\_match from University of Washington Genome Center PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center: <http://www.genome.washington.edu/LowQuality> Sequence: 21 contiguous PHRED high quality bases followed by vector sequence. Very Low Quality Sequence: Trace file contained 21 contiguous distinct peaks following vector sequence. Polyadenylation: Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this cDNA insert was polyadenylated.

Plate: LICM5 row: L column: 10

High quality sequence stop: 21.

Location/Qualifiers

1..21

/organism="Homo sapiens"

FEATURES

source

/db\_xref="taxon:9606"  
/clone="IMAGE:2821017"  
/clone\_lib="NIH\_MGC\_7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"

/notes="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 2 a 0 c 0 g 19 t  
ORIGIN

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x AW248782/rev ..

Align seg 1/1 to reverse of: AW248782 from: 1 to: 21

1 LysLysAsn 3  
|||||  
20 AAAAAAAT 12

seq\_name: gb\_est2:TI7609

seq\_documentation\_block:

LOCUS TI7609 21 bp mRNA linear EST 06-JUN-1994  
DEFINITION mps v321 The blue guys library Saccharomyces cerevisiae cDNA sequence upstream of LacZ fusion, mRNA sequence.

ACCESSION TI7609

VERSION TI7609.1 GI:458631

KEYWORDS EST.

SOURCE baker's yeast.

ORGANISM Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomyces.

1 (bases 1 to 21)

Burns,N., Grimwade,B., Ross-Macdonald,P.B., Choi,E.-Y., Finberg,K.,

Roeder,G.S. and Snyder,M.

Large-scale analysis of gene expression, protein localization and gene disruption in Saccharomyces cerevisiae

Genes Dev. 8, 1087-1105 (1994)

95011603

Contact: Snyder M

Department of Biology

Yale University

New Haven CT 06520-8103

Tel: 2034326139

Fax: 2034326161

Email: [snymicp@yalevm.ycc.yale.edu](mailto:snymicp@yalevm.ycc.yale.edu)

LacZ fusion; Vegetative expression; Beta-gal fusion localization pattern:

50 cytoplasmic spots. Sequence below near or adjacent to lacZ.

Seq primer: LacZ sequences in transposon.

Location/Qualifiers

1..21

/organism="Saccharomyces cerevisiae"

/db\_xref="taxon:4932"

/clone\_lib="The blue guys library"

/lab\_host="E.coli"

/notes="Vector: pRECtn; A yeast genomic DNA library was prepared in the vector pHS56, and subjected to transposon mutagenesis with mtn3. This mini-transposon carries lacZ sequences that lack an initiation codon; expression of



lacZ is only provided by in frame fusion to yeast coding sequence. The yeast genomic DNA carrying the transposon was excised from pHSG6 and transplanted back onto the yeast chromosome. Yeast colonies expressing lacZ were screened for in a color assay. A plasmid containing the genomic DNA/lacZ fusion junction was recovered from each individual yeast colony that expressed lacZ activity. These recovered plasmids comprise 'The blue guys library'. The fusion junction was then sequenced to identify the expressed ORF upstream of the fusion."

```

BASE COUNT      5 a      4 c      1 g      11 t
ORIGIN
alignment_scores:
  Quality:    16.00      Length:    3
  Ratio:      5.333      Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-2 x T17609/rev ..
Align seg 1/1 to reverse of: T17609 from: 1 to: 21

```

```

1 LysLysAsn 3
|||||
21 AAAAAAAT 13

```

```
seq_name: gb_gss:AZ308846
```

```

seq_documentation_block:
LOCUS      AZ308846
DEFINITION 1M0012H15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0012H15 F, DNA sequence.
ACCESSION  AZ308846
VERSION     AZ308846.1 GI:10349246
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus

```

```

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0012 row: H column: 15
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1..21
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0012H15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson

```

```

FEATURES
source

```

Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT      0 a      2 c      1 g      18 t
ORIGIN

```

```

alignment_scores:
  Quality:    16.00      Length:    3
  Ratio:      5.333      Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-2 x AZ308846/rev ..

```

```
Align seg 1/1 to reverse of: AZ308846 from: 1 to: 21
```

```

1 LysLysAsn 3
|||||
17 AAAAAAAC 9

```

```
seq_name: gb_gss:AZ309774
```

```

seq_documentation_block:
LOCUS      AZ309774
DEFINITION 1M0017F01F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0017F01 F, DNA sequence.
ACCESSION  AZ309774
VERSION     AZ309774.1 GI:10351103
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0017 row: F column: 01
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1..21
/organism="Mus musculus"
/strain="C57BL/6J"

```

```

FEATURES
source

```



Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 21.  
FEATURES  
    source  
        1..21  
            /organism="Mus musculus"  
            /strain="C57BL/6J"  
            /db\_xref="taxon:10090"  
            /clone\_lib="UUGC1M0033C05"  
            /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
            /sex="Male"  
            /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
            /note="Vector: PWD42nv; Purified genomic DNA from M.  
                musculus C57BL/6J (male) was obtained from the Jackson  
                Laboratory Mouse DNA Resource  
                (http://www.jax.org/resources/documents/dnares/). The DNA  
                was hydrodynamically sheared by repeated passage through a  
                0.005 inch orifice at constant velocity. The sheared DNA  
                was blunt end-repaired with T4 DNA polymerase and T4  
                polynucleotide kinase. Adaptor oligonucleotides were  
                ligated to the blunt ends in high molar excess. The  
                adaptored DNA was purified and size-selected for a 9.5 to  
                10.5 kb range using preparative agarose gel  
                electrophoresis. Vector DNA was prepared from a derivative  
                of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
                inducible derivative of plasmid R1. The vector was ligated  
                with adaptors complementary to the insert adaptors and  
                purified. The sheared, adaptored mouse DNA was annealed to  
                adaptored vector DNA, and transformed into  
                chemically-competent E. coli XL10-Gold (Stratagene) cells  
                and selected for ampicillin resistance."

BASE COUNT 18 a 3 c 0 g 0 t  
ORIGIN  
alignment\_scores:  
    Quality: 16.00 Length: 3  
    Ratio: 5.333 Gaps: 0  
    Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-528-682-2 x AZ316019 ..  
Align seg 1/1 to: AZ316019 from: 1 to: 21  
1 LysLysAsn 3  
|||||  
9 AAAAAAAC 17

/db\_xref="taxon:10090"  
/clone="UUGC1M0017F01"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 0 a 0 c 11 g 10 t  
ORIGIN  
alignment\_scores:  
    Quality: 16.00 Length: 3  
    Ratio: 5.333 Gaps: 0  
    Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-528-682-2 x AZ309774/rev ..  
Align seg 1/1 to reverse of: AZ309774 from: 1 to: 21  
1 LysLysAsn 3  
|||||  
19 AAAAAAAC 11  
seq\_name: gb\_gss:AZ316019  
seq\_documentation\_block:  
LOCUS AZ316019 21 bp DNA linear GSS 29-SEP-2000  
DEFINITION IM0033C05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0033C05 R, DNA sequence.  
ACCESSION AZ316019  
VERSION AZ316019.1 GI:10363426  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 21)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: gdunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0033 row: C column: 05

seq\_name: gb\_gss:AZ316019  
seq\_documentation\_block:  
LOCUS AZ316019 21 bp DNA linear GSS 29-SEP-2000  
DEFINITION IM0033C05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0033C05 R, DNA sequence.  
ACCESSION AZ316019  
VERSION AZ316019.1 GI:10363426  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 21)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: gdunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0033 row: C column: 05











OM of: US-09-528-682-2 to: Issued\_Patents\_NA:\* out\_format : pfs

Date: Jun 18, 2002 7:40 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=Cgcn2\_1/USPTO.spool/US09528682/runat\_18062002\_082443\_7827/app\_query.fasta\_1.689  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=p2n.rni  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682\_ECGN1\_1\_96 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-2

Query length: 3

Database: Issued\_Patents\_NA:\*

Database sequences: 383533

Database length: 122816752

Search time (sec): 143.210000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-994-824-3 +	16.00	103.50	446.97	10	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-994-824-3 +	16.00	103.50	446.97	10	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-404-430-4 -	16.00	103.50	446.97	10	1
/cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq:PCT-US96-09915-2 +	16.00	103.50	446.97	10	1
/cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq:PCT-US96-09915-3 +	16.00	103.50	446.97	10	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-794-9 +	16.00	102.16	530.83	12	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-007-005-9 +	16.00	102.16	530.83	12	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-247-390-9 +	16.00	102.16	530.83	12	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-796-9 +	16.00	102.16	530.83	12	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-393-783A-39 +	16.00	102.16	530.83	12	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-621-502A-3 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-781-550-19 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-544-381B-198 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-308-367-9 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-380-367-10 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-380-638A-11 -	16.00	101.57	572.45	13	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-778-794A-7 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-424-921-5 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-424-921-6 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-455-627-8 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-689-855-8 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-556-355A-5 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-556-355A-6 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-07-803-627A-5 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-07-803-627A-6 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-849-021-11 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-765-340-164 -	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-787-321-8 +	16.00	101.02	613.88	14	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-355 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-356 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-357 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-358 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-359 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-292-620A-360 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-29 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-30 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-31 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-32 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-33 -	16.00	100.52	655.15	15	1

/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-3 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-956-182-32 +	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-355 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-356 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-357 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-358 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-359 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-071-845-360 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-038-073-29 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-038-073-30 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-038-073-31 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-038-073-32 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-038-073-33 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-082-649B-69 -	16.00	100.52	655.15	15	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-703-778D-14 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-993-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-244-993-12 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-284-484A-4 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-426-807-3 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-861-306-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-861-306-12 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-468-037A-32 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-973A-32 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-1 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-2 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-3 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-4 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-5 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-6 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-7 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-8 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-435-434-5 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-016-520-37 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-016-520-38 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-016-520-39 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-016-520-40 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-016-520-41 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-144-611-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-378-568-1 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-3 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-9 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-10 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-12 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-30 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-130-973-47 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-3 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-9 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-10 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-12 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-30 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-477-902-47 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-123-572-1 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-564-399A-1 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-564-399A-2 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-564-399A-3 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-564-399A-4 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-378-655A-24 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-453-514A-11 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-202-294-3 -	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-390-850-597 +	16.00	100.04	696.27	16	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-790 +	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1068 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1069 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1070 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1071 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1072 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1073 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1074 -	16.00	99.60	737.23	17	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-373-124A-1075 -	16.00	99.60	737.23	17	1







/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-183-650-23 -	16.00	98.04	899.81	21	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:PCT-US92-03222-24 +	16.00	97.37	980.41	2	2
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-228-942-8 -	16.00	98.04	899.81	21	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-741-940-55 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-726-278-2 +	16.00	98.04	899.81	21	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-4 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-085-476-3 -	16.00	98.04	899.81	21	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-5 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-054-404-11 -	16.00	98.04	899.81	21	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-6 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-052-404-11 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-9 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-474-542A-148 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-10 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-474-542A-149 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-014-943A-11 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-123-449A-1 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-478-470-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-123-449A-2 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-214-559-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-648-148 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-473-015-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-648-149 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-289-548A-55 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-479-156-11 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-480-220A-6 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-480-220A-5 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-43 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-050-1 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-44 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-050-2 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-45 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-690-300-1 -	16.00	97.70	940.16	22	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-46 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-864-404-5 -	16.00	97.70	940.16	22	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-47 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-134-566-11 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-421-48 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-134-566-12 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-452-654-55 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-474-700B-20 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-477-559-11 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-235-936C-15 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-368-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-950-196-1 -	16.00	97.70	940.16	22	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-43 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-950-196-2 -	16.00	97.70	940.16	22	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-44 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-802-468-7 +	16.00	97.70	940.16	22	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-45 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-433-322B-3 +	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-46 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:PCT-US95-05812-20 -	16.00	97.70	940.16	22	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-47 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-872-644-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-911-48 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-977-284A-106 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-452-655B-55 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-494-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-477-306-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-018-129-15 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-43 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-510-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-44 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-621-914A-5 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-45 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-621-914A-6 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-46 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-619-301A-5 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-47 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-869-933-7 +	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-809-48 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-537-002A-17 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-995-161-11 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-479-532-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-864-04A-6 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-455-526-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-700-448-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-455-525-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-983-108-1 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-674-168-23 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-063-927-11 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-007-107-6 -	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-728D-32 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-639-080-26 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-835-728D-136 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-378-617A-6 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-118-841B-8 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-256-426B-106 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-362-535-25 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-761-277A-78 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-004-113-24 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-632-470-13 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-196-857-3 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-955-138-44 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-450-582-55 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-448-250-15 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-923-386A-13 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-929-922B-5 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-414-439-8 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-929-922B-14 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-461-637-262 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-985-908-9 -	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-796-10 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-804-439A-26 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-186-170-15 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-139-491-24 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-007-005-10 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-720-229-26 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-247-130-10 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-105-454-8 -	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-490-558-32 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-863-010-17 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-490-558-136 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-056-052-8 -	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-461-637-262 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-056-052-9 -	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-796-10 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-730-22 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-906-156A-82 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-342-394-5 +	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-434-131A-6 -	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-342-394-14 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-434-131A-10 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-024-429-17 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-686-933A-25 +	16.00	97.06	1.0e+03	24	1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-103-663-7 +	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-126-564A-24 -	16.00	96.76	1.1e+03	25	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-817-913-25 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-161-281A-10 -	16.00	96.76	1.1e+03	25	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-580-064-5 +	16.00	97.37	980.41	23	1	Se	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-478-675-61 -	16.00	96.76	1.1e+03	25	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-580-064-14 +	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-478-675-61 -	16.00	96.76	1.1e+03	25	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-371-257A-6 -	16.00	97.37	980.41	23	1	S	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-424-268-18 +	16.00	96.76	1.1e+03	25	1
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-985-916-20 -	16.00	97.37	980.41	23	1	S	/					



















```
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-202-206-1 + 16.00 93.12 1.7e+03 41 1 Se
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-06997-23 - 16.00 93.12 1.7e+03 41 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-908-317-33 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-250-859-3 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-490-803-3 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-758-626-21 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-443-568B-22 + 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-453-956-7 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-086-631-7 - 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-629-600-13 + 16.00 92.95 1.7e+03 42 1 Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-741-881-3 - 16.00 92.95 1.7e+03 42 1 Se

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-994-824-2
seq_documentation_block:
; Sequence 2, Application US/08994824
; Patent No. 6008400
; GENERAL INFORMATION:
; APPLICANT: Scarlinge, Stephen A.
; APPLICANT: Caruthers, Marvin H.
; TITLE OF INVENTION: NOVEL PROTECTING GROUPS AND USE THEREOF
; TITLE OF INVENTION: IN AN IMPROVED PROCESS FOR OLIGONUCLEOTIDE SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dan Cleveland, Jr.
; STREET: 1790 30th Street, Suite 140
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,824
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/488,878
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cleveland, Dan
; REGISTRATION NUMBER: 36,106
; REFERENCE/DOCKET NUMBER: 9028/103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)449-9497
; TELEFAX: (303)449-0814
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; US-08-994-824-2

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-994-824-2 ..
Align seg 1/1 to: US-08-994-824-2 from: 1 to: 10

1 LysLysAsn 3
|||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-404-430-4
seq_documentation_block:
; Sequence 4, Application US/09404430
; Patent No. 6350853
; GENERAL INFORMATION:
; APPLICANT: Neilsen, Peter
; APPLICANT: Knudsen, Helle
; TITLE OF INVENTION: Conjugated Peptide Nucleic Acids Having Enhanced
; TITLE OF INVENTION: Cellular Uptake
; FILE REFERENCE: ISIS4181
; CURRENT APPLICATION NUMBER: US/09/404,430
```



; CURRENT FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US08/564,765  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: No. 6350853e1 Sequence  
US-09-404-430-4

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-404-430-4/rev ..

Align seg 1/1 to reverse of: US-09-404-430-4 from: 1 to: 10

1 LysLysAsn 3  
|||||  
10 AAAAAAAT 2

seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:PCT-US96-09915-2

seq\_documentation\_block:  
; Sequence 2, Application PC/TUS9609915  
; GENERAL INFORMATION:  
; APPLICANT: Scarlinge, Stephen A.  
; APPLICANT: Caruthers, Marvin H.  
; TITLE OF INVENTION: NOVEL PROTECTING GROUPS AND USE THEREOF  
; TITLE OF INVENTION: IN AN IMPROVED PROCESS FOR OLIGONUCLEOTIDE SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steven C. Petersen  
; STREET: 1900 Fifteenth Street  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80302  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09915  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/488,878  
; FILING DATE: 09-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Petersen, Steven C.  
; REGISTRATION NUMBER: 36,238  
; REFERENCE/DOCKET NUMBER: file no. 16840-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-546-1300  
; TELEFAX: 303-449-5426  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA (synthetic)"  
PCT-US96-09915-2

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x PCT-US96-09915-2 ..

Align seg 1/1 to: PCT-US96-09915-2 from: 1 to: 10

1 LysLysAsn 3  
|||||  
2 AAAAAAAT 10

seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:PCT-US96-09915-3

seq\_documentation\_block:

; Sequence 3, Application PC/TUS9609915  
; GENERAL INFORMATION:  
; APPLICANT: Scarlinge, Stephen A.  
; APPLICANT: Caruthers, Marvin H.  
; TITLE OF INVENTION: NOVEL PROTECTING GROUPS AND USE THEREOF  
; TITLE OF INVENTION: IN AN IMPROVED PROCESS FOR OLIGONUCLEOTIDE SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steven C. Petersen  
; STREET: 1900 Fifteenth Street  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80302  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09915  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/488,878  
; FILING DATE: 09-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Petersen, Steven C.  
; REGISTRATION NUMBER: 36,238  
; REFERENCE/DOCKET NUMBER: file no. 16840-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-546-1300  
; TELEFAX: 303-449-5426  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "RNA (synthetic)"  
PCT-US96-09915-3

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x PCT-US96-09915-3 ..

Align seg 1/1 to: PCT-US96-09915-3 from: 1 to: 10



```
1 LysLysAsn 3
|||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-794A-9

seq_documentation_block:
; Sequence 9, Application US/09244794A
; Patent No. 6214553
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350006
; CURRENT APPLICATION NUMBER: US/09/244,794A
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/035,963
; PRIOR FILING DATE: 1997-01-27
; PRIOR APPLICATION NUMBER: 60/064,491
; PRIOR FILING DATE: 1997-11-06
; PRIOR APPLICATION NUMBER: 09/007,005
; PRIOR FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
US-09-244-794A-9

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-244-794A-9 ..
Align seg 1/1 to: US-09-244-794A-9 from: 1 to: 12

1 LysLysAsn 3
|||||
3 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-247-190-9

seq_documentation_block:
; Sequence 9, Application US/09247190
; Patent No. 6261804
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350005
; CURRENT APPLICATION NUMBER: US/09/247,190
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
US-09-247-190-9

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-247-190-9 ..
Align seg 1/1 to: US-09-247-190-9 from: 1 to: 12

1 LysLysAsn 3
|||||
3 AAAAAAAT 11

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-244-796-9

seq_documentation_block:
; Sequence 9, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
```



;; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
;; FILE REFERENCE: 00786/350007  
;; CURRENT APPLICATION NUMBER: US/09/244,796  
;; CURRENT FILING DATE: 1999-02-05  
;; EARLIER APPLICATION NUMBER: 60/035,963  
;; EARLIER FILING DATE: 1997-01-27  
;; EARLIER APPLICATION NUMBER: 60/064,491  
;; EARLIER FILING DATE: 1997-11-06  
;; EARLIER APPLICATION NUMBER: 09/007,005  
;; EARLIER FILING DATE: 1998-01-14  
;; NUMBER OF SEQ ID NOS: 33  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 9  
;; LENGTH: 12  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Translation template  
US-09-244-796-9

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-244-796-9 ..

Align seg 1/1 to: US-09-244-796-9 from: 1 to: 12

1 LysLysasn 3  
|||||  
3 AAAAAAAC 11

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-393-783A-39

seq\_documentation\_block:  
; Sequence 39, Application US/09393783A  
; Patent No. 6355428  
; GENERAL INFORMATION:  
; APPLICANT: Schroth, Gary P.  
; APPLICANT: Bruice, Thomas Wayne  
; APPLICANT: Sub. Young J.  
; TITLE OF INVENTION: Nucleic Acid Ligand Interaction Assays  
; FILE REFERENCE: 4600-0128.30  
; CURRENT APPLICATION NUMBER: US/09/393.783A  
; CURRENT FILING DATE: 1999-10-09  
; PRIOR APPLICATION NUMBER: US 09/151,890  
; PRIOR FILING DATE: 1998-09-11  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 12  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_binding  
; LOCATION: (1)...(12)  
; OTHER INFORMATION: synthesized test oligonucleotide for binding  
; OTHER INFORMATION: studies  
US-09-393-783A-39

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-393-783A-39 ..

Align seg 1/1 to: US-09-393-783A-39 from: 1 to: 12

1 LysLysasn 3  
|||||  
3 AAAAAAAC 11

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-621-502A-3

seq\_documentation\_block:  
; Sequence 3, Application US/08621502A  
; Patent No. 5763591  
; GENERAL INFORMATION:  
; APPLICANT: Toyoda Ph.D., Hiroo  
; APPLICANT: Formby Ph.D., Bent  
; TITLE OF INVENTION: POLYNUCLEIC ACID SEQUENCES THAT ARE  
; FUNCTIONALLY ASSOCIATED WITH THE DEVELOPMENT OF AUTOIMMUNE  
; DISEASE  
; TITLE OF INVENTION: DISEASE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
; STREET: 444 South Flower Street, Suite 2000  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90071

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/621,502A  
; FILING DATE: 25-MAR-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fujita Esq., Sharon M.  
; REGISTRATION NUMBER: 38,459  
; REFERENCE/DOCKET NUMBER: P07 34324  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 213/622-7700  
; TELEFAX: 213/489-4210  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-08-621-502A-3

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-08-621-502A-3/rev ..

Align seg 1/1 to reverse of: US-08-621-502A-3 from: 1 to: 13

1 LysLysasn 3  
|||||  
10 AAAAAAAC 2

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-781-550-19

seq\_documentation\_block:  
; Sequence 19, Application US/08781550  
; Patent No. 5861242  
; GENERAL INFORMATION:  
; APPLICANT: Chee, Mark



```
; APPLICANT: Gingeras, Thomas R.
; APPLICANT: Fodor, Stephen P.A.
; APPLICANT: Hubbell, Earl A.
; APPLICANT: Morris, Macdonald S.
; TITLE OF INVENTION: HIV Diagnosis by Arrays of Nucleic Acid
; TITLE OF INVENTION: Probes on Biological Chips
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,550
; FILING DATE: 09-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,064
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: US 08/143,312
; FILING DATE: 26-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/082,937
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1046.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
; US-08-781-550-19

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-08-781-550-19/rev ..
Align seg 1/1 to reverse of: US-08-781-550-19 from: 1 to: 13

  1 LysLysAsn 3
  |||
  13 AAAAAAAC 5

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-544-381B-198
seq_documentation_block:
; Sequence 198, Application US/08544381B
; Patent No. 6027880
; GENERAL INFORMATION:
; APPLICANT: Cronin, Maureen T.
; APPLICANT: Miyada, Charles Garrett
; APPLICANT: Hubbell, Earl A.
; APPLICANT: Chee, Mark
; APPLICANT: Fodor, Stephen P.A.
```

```
; APPLICANT: Huang, Xiaohua C.
; APPLICANT: Lipshutz, Robert J.
; APPLICANT: Lobban, Peter E.
; APPLICANT: Morris, Macdonald S.
; APPLICANT: Sheldon, Edward L.
; TITLE OF INVENTION: Arrays of Nucleic Acid Probes for
; TITLE OF INVENTION: Detecting Cystic Fibrosis
; NUMBER OF SEQUENCES: 250
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,381B
; FILING DATE: 10-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/510,521
; FILING DATE: 02-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12305
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,064
; FILING DATE: 02-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/143,312
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018547-0041300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 198:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (probe)
; US-08-544-381B-198

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-08-544-381B-198/rev ..
Align seg 1/1 to reverse of: US-08-544-381B-198 from: 1 to: 13

  1 LysLysAsn 3
  |||
  13 AAAAAAAC 5

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-308-367-9
seq_documentation_block:
; Sequence 9, Application US/09308367A
; Patent No. 6043060
; GENERAL INFORMATION:
```



; APPLICANT: IMANISHI, Takeshi  
; TITLE OF INVENTION: NOVEL NUCLEOTIDE ANALOGUES  
; FILE REFERENCE: IMANISHI-1  
; CURRENT APPLICATION NUMBER: US/09/308,367A  
; CURRENT FILING DATE: 1999-05-18  
; EARLIER APPLICATION NUMBER: PCT/JP97/04187  
; EARLIER FILING DATE: 1997-11-18  
; EARLIER APPLICATION NUMBER: JAPAN 306585/1996  
; EARLIER FILING DATE: 1996-11-18  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 13  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-367-9

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-308-367-9/rev ..

Align seg 1/1 to reverse of: US-09-308-367-9 from: 1 to: 13

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-308-367-10

seq\_documentation\_block:  
; Sequence 10, Application US/09308367A  
; Patent No. 6043060  
; GENERAL INFORMATION:  
; APPLICANT: IMANISHI, Takeshi  
; TITLE OF INVENTION: NOVEL NUCLEOTIDE ANALOGUES  
; FILE REFERENCE: IMANISHI-1  
; CURRENT APPLICATION NUMBER: US/09/308,367A  
; CURRENT FILING DATE: 1999-05-18  
; EARLIER APPLICATION NUMBER: PCT/JP97/04187  
; EARLIER FILING DATE: 1997-11-18  
; EARLIER APPLICATION NUMBER: JAPAN 306585/1996  
; EARLIER FILING DATE: 1996-11-18  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 13  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; FEATURE:  
; OTHER INFORMATION: n is 3'-O, 4'-C-methylene uridine  
US-09-308-367-10

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-308-367-10/rev ..

Align seg 1/1 to reverse of: US-09-308-367-10 from: 1 to: 13

1 LysLysAsn 3

|||||  
9 AAAAAAAC 1  
seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-380-638A-10

seq\_documentation\_block:  
; Sequence 10, Application US/09380638A  
; Patent No. 6268490  
; GENERAL INFORMATION:  
; APPLICANT: IMANISHI, Takeshi  
; APPLICANT: OBIKA, Satoshi  
; TITLE OF INVENTION: NOVEL BICYCLONUCLEOSIDE AND OLIGONUCLEOTIDE ANALOGUE  
; FILE REFERENCE: IMANISHI-2  
; CURRENT APPLICATION NUMBER: US/09/380,638A  
; CURRENT FILING DATE: 1999-09-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/00945  
; PRIOR FILING DATE: 1998-03-09  
; PRIOR APPLICATION NUMBER: JP 53409/1997  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 13  
; TYPE: DNA  
; ORGANISM: synthetic construct  
US-09-380-638A-10

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-380-638A-10/rev ..

Align seg 1/1 to reverse of: US-09-380-638A-10 from: 1 to: 13

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-380-638A-11

seq\_documentation\_block:  
; Sequence 11, Application US/09380638A  
; Patent No. 6268490  
; GENERAL INFORMATION:  
; APPLICANT: IMANISHI, Takeshi  
; APPLICANT: OBIKA, Satoshi  
; TITLE OF INVENTION: NOVEL BICYCLONUCLEOSIDE AND OLIGONUCLEOTIDE ANALOGUE  
; FILE REFERENCE: IMANISHI-2  
; CURRENT APPLICATION NUMBER: US/09/380,638A  
; CURRENT FILING DATE: 1999-09-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/00945  
; PRIOR FILING DATE: 1998-03-09  
; PRIOR APPLICATION NUMBER: JP 53409/1997  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 13  
; TYPE: DNA  
; ORGANISM: synthetic construct  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: n at positions 11 and 12 is unknown  
US-09-380-638A-11

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0



Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-380-638A-11/rev ..

Align seg 1/1 to reverse of: US-09-380-638A-11 from: 1 to: 13

1 LysLysAsn 3  
|||||  
9 AAAAAAAC 1

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-778-794A-7

## seq\_documentation\_block:

; Sequence 7, Application US/08778794A

; Patent No. 6309823

## ; GENERAL INFORMATION:

; APPLICANT: Cronin, Maureen T.

; APPLICANT: Miyada, Charles Garrett

; APPLICANT: Hubbell, Earl A.

; APPLICANT: Chee, Mark

; APPLICANT: Fodor, Stephen P.A.

; APPLICANT: Huang, Xiaohua C.

; APPLICANT: Lipshutz, Robert J.

; APPLICANT: Lobban, Peter E.

; APPLICANT: Morris, MacDonald S.

; APPLICANT: Sheldon, Edward L.

; TITLE OF INVENTION: Arrays of Nucleic Acid Probes

; TITLE OF INVENTION: for Analyzing Biotransformation Genes

; NUMBER OF SEQUENCES: 156

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/778,794A

; FILING DATE: 03-JAN-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/143,312

; FILING DATE: 26-OCT-1993

; APPLICATION NUMBER: US 08/284,064

; FILING DATE: 02-AUG-1994

; APPLICATION NUMBER: WO PCT/US94/12305

; FILING DATE: 26-OCT-1994

; APPLICATION NUMBER: US 08/510,521

; FILING DATE: 02-AUG-1995

; APPLICATION NUMBER: US 08/544,381

; FILING DATE: 10-OCT-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Liebeschuetz, Joe

; REGISTRATION NUMBER: 37,505

; REFERENCE/DOCKET NUMBER: 018547-015700US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0200

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; US-08-778-794A-7

## alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-08-778-794A-7/rev ..

Align seg 1/1 to reverse of: US-08-778-794A-7 from: 1 to: 13

1 LysLysAsn 3

|||||

13 AAAAAAAC 5

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-424-921-5

## seq\_documentation\_block:

; Sequence 5, Application US/08424921

; Patent No. 5545552

; GENERAL INFORMATION:

; APPLICANT: Machur, Eric A

; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCCLUS

; TITLE OF INVENTION: FURIOSUS DNA POLYMERASE I

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bingham & Fitting

; STREET: 12526 High Bluff Road, Suite 300

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92130

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/424,921

; FILING DATE: 19-APR-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/803,627

; FILING DATE: 02-DEC-1991

; APPLICATION NUMBER: US 07/620,568

; FILING DATE: 03-DEC-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/657,073

; FILING DATE: 19-FEB-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/776,552

; FILING DATE: 15-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: STG0100P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-792-3680

; TELEFAX: 619-792-8477

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-424-921-5

## alignment\_scores:



Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-08-424-921-5 ..

Align seg 1/1 to: US-08-424-921-5 from: 1 to: 14

1 LysLysAsn 3

|||||

4 AAGAAAGAAC 12

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-424-921-6

## seq\_documentation\_block:

; Sequence 6, Application US/08424921

; Patent No. 5545552

; GENERAL INFORMATION:

; APPLICANT: Mathur, Eric A

; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCCLUS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bingham & Fitting

; STREET: 12526 High Bluff Road, Suite 300

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92130

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/424,921

; FILING DATE: 19-APR-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/803,627

; FILING DATE: 02-DEC-1991

; APPLICATION NUMBER: US 07/620,568

; FILING DATE: 03-DEC-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/657,073

; FILING DATE: 19-FEB-1991

; APPLICATION NUMBER: US 07/776,552

; FILING DATE: 15-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: STG0100P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-792-3680

; TELEFAX: 619-792-8477

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-424-921-6

## alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-08-424-921-6/rev ..

Align seg 1/1 to reverse of: US-08-424-921-6 from: 1 to: 14

1 LysLysAsn 3

|||||

11 AAGAAAGAAC 3

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-455-627-8

## seq\_documentation\_block:

; Sequence 8, Application US/08455627

; Patent No. 5571677

; GENERAL INFORMATION:

; APPLICANT: Sergei M. Gryaznov

; TITLE OF INVENTION: Convergent Synthesis of Branched and Multiply

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooley Godward LLP

; STREET: Five Palo Alto Square, 3000 El Camino Real

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94306-2155

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,627

; FILING DATE: 31-MAY-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Nakamura, Jackie N.

; REGISTRATION NUMBER: 35,966

; REFERENCE/DOCKET NUMBER: LYNX-003/01 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-843-5000

; TELEFAX: 415-857-0663

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

US-08-455-627-8

## alignment\_scores:

Quality: 16.00 Length: 3

Ratio: 5.333 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-08-455-627-8 ..

Align seg 1/1 to: US-08-455-627-8 from: 1 to: 14

1 LysLysAsn 3

|||||

6 AAAAAAAC 14

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-689-856-8

## seq\_documentation\_block:

; Sequence 8, Application US/08689856

; Patent No. 5830658

; GENERAL INFORMATION:



APPLICANT: Sergei M. Gryaznov  
TITLE OF INVENTION: Convergent Synthesis of Branched and Multiply  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward LLP  
STREET: Five Palo Alto Square, 3000 El Camino Real  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306-2155  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/689,856  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,627  
FILING DATE: 31-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakamura, Jackie N.  
REGISTRATION NUMBER: 35,966  
REFERENCE/DOCKET NUMBER: LYNX-003/01 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5000  
TELEFAX: 415-857-0663  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-689-856-8

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-689-856-8 ..

Align seg 1/1 to: US-08-689-856-8 from: 1 to: 14

1 LysLysAsn 3  
|||||||  
6 AAAAAAAC 14

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-556-355A-5

seq\_documentation\_block:  
Sequence 5, Application US/08556355A  
Patent No. 5866395  
GENERAL INFORMATION:  
APPLICANT: MATHUR, Eric J.  
TITLE OF INVENTION: Purified Thermostable Pyrococcus  
TITLE OF INVENTION: Purified Thermostable Pyrococcus  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/556,355A  
FILING DATE: 13-NOV-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/424,921  
FILING DATE: 19-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/803,627  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,846  
FILING DATE: 21-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/657,073  
FILING DATE: 19-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,568  
FILING DATE: 03-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Barker, M. Paul  
REGISTRATION NUMBER: 32,013  
REFERENCE/DOCKET NUMBER: 04121.0004-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
US-08-556-355A-5

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-556-355A-5 ..

Align seg 1/1 to: US-08-556-355A-5 from: 1 to: 14

1 LysLysAsn 3  
|||||||  
4 AARARAA 12

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-556-355A-6

seq\_documentation\_block:  
Sequence 6, Application US/08556355A  
Patent No. 5866395  
GENERAL INFORMATION:  
APPLICANT: MATHUR, Eric J.  
TITLE OF INVENTION: Purified Thermostable Pyrococcus  
TITLE OF INVENTION: Purified Thermostable Pyrococcus  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:



```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556.355A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/424,921
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,627
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,846
; FILING DATE: 21-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-556-355A-6

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-556-355A-6/rev ..

Align seg 1/1 to reverse of: US-08-556-355A-6 from: 1 to: 14
```

```
1 LysLysAsn 3
|||||
11 AAAAAAAAY 3

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-07-803-627A-5

seq_documentation_block:
; Sequence 5, Application US/07803627A
; Patent No. 5948663
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803.627A
; FILING DATE: 02-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,552
; FILING DATE: 14-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-07-803-627A-5

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-07-803-627A-5 ..

Align seg 1/1 to: US-07-803-627A-5 from: 1 to: 14

1 LysLysAsn 3
|||||
4 AAAAAAAAY 12

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-07-803-627A-6

seq_documentation_block:
; Sequence 6, Application US/07803627A
; Patent No. 5948663
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,627A
```



```
; FILING DATE: 02-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,552
; FILING DATE: 14-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; US-07-803-627A-6

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-07-803-627A-6/rev ..
Align seg 1/1 to reverse of: US-07-803-627A-6 from: 1 to: 14

1 LysLysAsn 3
|||||
11 AARARAAY 3

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-849-021-11

seq_documentation_block:
; Sequence 11, Application US/08849021
; Patent No. 5953276
; GENERAL INFORMATION:
; APPLICANT: MORGANTE, MICHELE
; APPLICANT: VOGEL, JULIE M.
; TITLE OF INVENTION: COMPOUND MICROSATELLITE
; TITLE OF INVENTION: PRIMERS FOR THE
; TITLE OF INVENTION: DETECTION OF GENETIC
; TITLE OF INVENTION: POLYMORPHISMS
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND
; ADDRESSEE: COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENT IN RELEASE #1.0, VERSION 1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,021
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/346,456
; FILING DATE: 28 NOVEMBER 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1064-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-992-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-849-021-11

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-849-021-11/rev ..
Align seg 1/1 to reverse of: US-08-849-021-11 from: 1 to: 14

1 LysLysAsn 3
|||||
12 AAAAAAAC 4

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-765-340-164

seq_documentation_block:
; Sequence 164, Application US/08765340
; Patent No. 6150092
; GENERAL INFORMATION:
; APPLICANT: UCHIDA, K.,
; APPLICANT: UCHIDA, T.,
; APPLICANT: TANAKA, Y.,
; APPLICANT: MATSUDA, Y.,
; APPLICANT: KONDO, S.
; TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID
; TITLE OF INVENTION: COMPOUND
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,340
; FILING DATE: 23-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 145146/94
; FILING DATE: 27-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 311130/94
; FILING DATE: 21-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SERUNIAN, LESLIE
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 1452-4005
```



```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 164:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; US-08-765-340-164

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-765-340-164/rev ..
Align seg 1/1 to reverse of: US-08-765-340-164 from: 1 to: 14

1 LysLysAsn 3
|||||
11 AAGAAAAAT 3

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-787-321-8

seq_documentation_block:
; Sequence 8, Application US/08787321A
; Patent No. 6180777
; GENERAL INFORMATION:
; APPLICANT: Horn, Thomas
; TITLE OF INVENTION: SYNTHESIS OF BRANCHED NUCLEIC ACIDS
; FILE REFERENCE: (1300)-1199.002
; CURRENT APPLICATION NUMBER: US/08/787,321A
; CURRENT FILING DATE: 1997-01-03
; EARLIER APPLICATION NUMBER: US PROV 60/009,918
; EARLIER FILING DATE: 1996-01-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-08-787-321-8

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-787-321-8 ..
Align seg 1/1 to: US-08-787-321-8 from: 1 to: 14

1 LysLysAsn 3
|||||
6 AAAAAAAC 14

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-292-620A-355

seq_documentation_block:
; Sequence 355, Application US/08292620A
; Patent No. 5837542
; GENERAL INFORMATION:
```

```
;
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,620A
; FILING DATE: August 17, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-292-620A-355

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-292-620A-355/rev ..
Align seg 1/1 to reverse of: US-08-292-620A-355 from: 1 to: 15

1 LysLysAsn 3
|||||
14 AAAAAAAT 6

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-292-620A-356

seq_documentation_block:
; Sequence 356, Application US/08292620A
; Patent No. 5837542
; GENERAL INFORMATION:
```

two



APPLICANT: Susan Grimm  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James McSwiggen  
APPLICANT: Sean Sullivan  
APPLICANT: Kenneth G. Draper  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
TITLE OF INVENTION: DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: INTRACELLULAR ADHESION  
TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)  
NUMBER OF SEQUENCES: 2390  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,620A  
FILING DATE: August 17, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/008,895  
FILING DATE: January 19, 1993  
APPLICATION NUMBER: 07/989,849  
FILING DATE: December 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 356:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-292-620A-356

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-292-620A-356/rev ..

Align seg 1/1 to reverse of: US-08-292-620A-356 from: 1 to: 15

1 LysLysasn 3  
|||||  
13 AAAAAAAT 5

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-292-620A-357

seq\_documentation\_block:  
; Sequence 357, Application US/08292620A  
; Patent No. 5837542  
; GENERAL INFORMATION:

APPLICANT: Susan Grimm  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James McSwiggen  
APPLICANT: Sean Sullivan  
APPLICANT: Kenneth G. Draper  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
TITLE OF INVENTION: DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: INTRACELLULAR ADHESION  
TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)  
NUMBER OF SEQUENCES: 2390  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,620A  
FILING DATE: August 17, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/008,895  
FILING DATE: January 19, 1993  
APPLICATION NUMBER: 07/989,849  
FILING DATE: December 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 357:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-292-620A-357

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-292-620A-357/rev ..

Align seg 1/1 to reverse of: US-08-292-620A-357 from: 1 to: 15

1 LysLysasn 3  
|||||  
12 AAAAAAAT 4

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-292-620A-358

seq\_documentation\_block:  
; Sequence 358, Application US/08292620A  
; Patent No. 5837542  
; GENERAL INFORMATION:

two

two



APPLICANT: Susan Grimm  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James McSwiggen  
APPLICANT: Sean Sullivan  
APPLICANT: Kenneth G. Draper  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
TITLE OF INVENTION: DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: INTRACELLULAR ADHESION  
TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)  
NUMBER OF SEQUENCES: 2390  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,620A  
FILING DATE: August 17, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/008,895  
FILING DATE: January 19, 1993  
APPLICATION NUMBER: 07/989,849  
FILING DATE: December 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 358:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-292-620A-358

two

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-292-620A-358/rev ..

Align seg 1/1 to reverse of: US-08-292-620A-358 from: 1 to: 15

1 LysLysAsn 3  
|||||||  
11 AAAAAAAT 3

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-292-620A-359

seq\_documentation\_block:  
; Sequence 359, Application US/08292620A  
; Patent No. 5837542  
; GENERAL INFORMATION:

APPLICANT: Susan Grimm  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James McSwiggen  
APPLICANT: Sean Sullivan  
APPLICANT: Kenneth G. Draper  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
TITLE OF INVENTION: DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: INTRACELLULAR ADHESION  
TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)  
NUMBER OF SEQUENCES: 2390  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,620A  
FILING DATE: August 17, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/008,895  
FILING DATE: January 19, 1993  
APPLICATION NUMBER: 07/989,849  
FILING DATE: December 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 359:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-292-620A-359

two

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-292-620A-359/rev ..

Align seg 1/1 to reverse of: US-08-292-620A-359 from: 1 to: 15

1 LysLysAsn 3  
|||||||  
10 AAAAAAAT 2

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-292-620A-360

seq\_documentation\_block:  
; Sequence 360, Application US/08292620A  
; Patent No. 5837542  
; GENERAL INFORMATION:



```
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,620A
; FILING DATE: August 17, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 360:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-292-620A-360

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-292-620A-360/rev ..
Align seg 1/1 to reverse of: US-08-292-620A-360 from: 1 to: 15

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-29
seq_documentation_block:
; Sequence 29, Application US/08585684B
; Patent No. 5877021
; GENERAL INFORMATION:
; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

two

; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: Jarvis, Thale
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/585,684B
; FILING DATE: January 16, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,951
; FILING DATE: July 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-585-684B-29

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-585-684B-29/rev ..
Align seg 1/1 to reverse of: US-08-585-684B-29 from: 1 to: 15

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-30
seq_documentation_block:
; Sequence 30, Application US/08585684B
; Patent No. 5877021
; GENERAL INFORMATION:
; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: Jarvis, Thale
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

1 LysLysasn 3
|||||||
14 AAGAAAAAT 6

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-585-684B-30
seq_documentation_block:
; Sequence 30, Application US/08585684B
; Patent No. 5877021
; GENERAL INFORMATION:
; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: Jarvis, Thale
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
```



; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,951  
; FILING DATE: January 16, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 218/078  
; FILING DATE: July 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 218/078  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-585-684B-30

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-585-684B-30/rev ..

Align seg 1/1 to reverse of: US-08-585-684B-30 from: 1 to: 15

1 LysLysAsn 3  
|||||  
13 AAGAAAAAT 5

seq\_name: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:US-08-585-684B-31

seq\_documentation\_block:  
; Sequence 31, Application US/08585684B  
; Patent No. 5877021  
; GENERAL INFORMATION:  
; APPLICANT: Stinchcomb, Daniel T.  
; APPLICANT: Jarvis, Thale  
; APPLICANT: McSwiggen, James  
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
; TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
; NUMBER OF SEQUENCES: 2751  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/585,684B  
; FILING DATE: January 16, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,951  
; FILING DATE: July 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 218/078  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-585-684B-31

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-08-585-684B-31/rev ..

Align seg 1/1 to reverse of: US-08-585-684B-31 from: 1 to: 15

1 LysLysAsn 3  
|||||  
12 AAGAAAAAT 4

seq\_name: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:US-08-585-684B-32

seq\_documentation\_block:  
; Sequence 32, Application US/08585684B  
; Patent No. 5877021  
; GENERAL INFORMATION:  
; APPLICANT: Stinchcomb, Daniel T.  
; APPLICANT: Jarvis, Thale  
; APPLICANT: McSwiggen, James  
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
; TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
; NUMBER OF SEQUENCES: 2751  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/585,684B  
; FILING DATE: January 16, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,951  
; FILING DATE: July 7, 1995



```
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-585-684B-32
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
  US-09-528-682-2 x US-08-585-684B-32/rev ..
```

Align seg 1/1 to reverse of: US-08-585-684B-32 from: 1 to: 15

```
1 LysLysasn 3
|||||||
11 AAGAAAAAT 3
```

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-585-684B-33

```
seq_documentation_block:
; Sequence 33, Application US/08585684B
; Patent No. 5877021
; GENERAL INFORMATION:
; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: Jarvis, Thale
; APPLICANT: McSwigen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
; TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/585,684B
; FILING DATE: January 16, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,951
; FILING DATE: July 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 33:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-585-684B-33

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-08-585-684B-33/rev ..

Align seg 1/1 to reverse of: US-08-585-684B-33 from: 1 to: 15

1 LysLysasn 3
|||||||
10 AAGAAAAAT 2

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-763-354-3

seq_documentation_block:
; Sequence 3, Application US/087633354
; Patent No. 5965721
; GENERAL INFORMATION:
; APPLICANT: Cook, Philip Dan
; APPLICANT: Sanghri, Yogesh H.
; APPLICANT: Vasseur, Jean J.
; APPLICANT: Debart, Francoise
; TITLE OF INVENTION: Backbone Modified Oligonucleotide Analogs
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/763,354
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 703,619
; FILING DATE: 21-MAY-1991
; APPLICATION NUMBER: US 566,836
; FILING DATE: 13-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 558,663
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lucci, Joseph
; REGISTRATION NUMBER: 33,307
; REFERENCE/DOCKET NUMBER: ISIS-2404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: NO
```



US-08-763-354-3

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-08-763-354-3/rev ..

Align seg 1/1 to reverse of: US-08-763-354-3 from: 1 to: 15

1 LysLysAsn 3

|||||

11 AAAAAAAC 3

seq\_name: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:US-08-956-182-32

seq\_documentation\_block:

; Sequence 32, Application US/08956182  
; Patent No. 6100450  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Terry L.  
; TITLE OF INVENTION: NOVEL SEED SPECIFIC PROMOTERS BASED ON  
; TITLE OF INVENTION: ARABIDOPSIS GENES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,182  
; FILING DATE:  
; CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10701  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-956-182-32

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-08-956-182-32 ..

Align seg 1/1 to: US-08-956-182-32 from: 1 to: 15

1 LysLysAsn 3

|||||

3 AAGAGAGAC 11

seq\_name: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:US-09-071-845-355

seq\_documentation\_block:

; Sequence 355, Application US/09071845  
; Patent No. 6132967  
; GENERAL INFORMATION:  
; APPLICANT: Susan Grimm  
; APPLICANT: Dan T. Stinchcomb  
; APPLICANT: James McSwiggen  
; APPLICANT: Sean Sullivan  
; APPLICANT: Kenneth G. Draper  
; TITLE OF INVENTION: RIBOZYME TREATMENT OF  
; TITLE OF INVENTION: DISEASES OR CONDITIONS  
; TITLE OF INVENTION: RELATED TO LEVELS OF  
; TITLE OF INVENTION: INTRACELLULAR ADHESION  
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)  
; NUMBER OF SEQUENCES: 2390  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066

COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,845  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/292,620  
; FILING DATE: August 17, 1994  
; APPLICATION NUMBER: 08/008,895  
; FILING DATE: January 19, 1993  
; APPLICATION NUMBER: 07/989,849  
; FILING DATE: December 7, 1992

ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 208/149  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 355:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-071-845-355

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-071-845-355/rev ..

Align seg 1/1 to reverse of: US-09-071-845-355 from: 1 to: 15

1 LysLysAsn 3

|||||



14 AAAAAAAT 6

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-071-845-356

```
seq_documentation_block:
; Sequence 356, Application US/09071845
; Patent No. 6132967
; GENERAL INFORMATION:
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,845
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,620
; FILING DATE: August 17, 1994
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 356:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-071-845-356
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-071-845-356/rev ..

Align seg 1/1 to reverse of: US-09-071-845-356 from: 1 to: 15

1 LysLysAsn 3

|||||

13 AAAAAAAT 5

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-071-845-357

```
seq_documentation_block:
; Sequence 357, Application US/09071845
; Patent No. 6132967
; GENERAL INFORMATION:
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,845
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,620
; FILING DATE: August 17, 1994
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 357:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-071-845-357
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-071-845-357/rev ..

Align seg 1/1 to reverse of: US-09-071-845-357 from: 1 to: 15

1 LysLysAsn 3

|||||



12 AAAAAAAAAAT 4

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMPB.seq:US-09-071-845-358

```
seq_documentation_block:
; Sequence 358, Application US/09071845
; Patent No. 6132967
; GENERAL INFORMATION:
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071.845
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/292.620
; FILING DATE: August 17, 1994
; APPLICATION NUMBER: 08/008.895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989.849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-071-845-358
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-071-845-358/rev ..

Align seg 1/1 to reverse of: US-09-071-845-358 from: 1 to: 15

1 LysLysAsn 3  
|||||

11 AAAAAAAAAAT 3

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMPB.seq:US-09-071-845-359

```
seq_documentation_block:
; Sequence 359, Application US/09071845
; Patent No. 6132967
; GENERAL INFORMATION:
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071.845
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/292.620
; FILING DATE: August 17, 1994
; APPLICATION NUMBER: 08/008.895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989.849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 359:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-071-845-359
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-071-845-359/rev ..

Align seg 1/1 to reverse of: US-09-071-845-359 from: 1 to: 15

1 LysLysAsn 3  
|||||



10 AAAAAAAT 2

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-071-845-360

```
seq_documentation_block:
; Sequence 360, Application US/09071845
; Patent No. 6132967
; GENERAL INFORMATION:
; APPLICANT: Susan Grimm
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: INTRACELLULAR ADHESION
; TITLE OF INVENTION: MOLECULE-1 (I-CAM-1)
; NUMBER OF SEQUENCES: 2390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,845
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,620
; FILING DATE: August 17, 1994
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 360:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-071-845-360
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-071-845-360/rev ..

Align seg 1/1 to reverse of: US-09-071-845-360 from: 1 to: 15

1 LysLysAsn 3

|||||

9 AAAAAAAT 1

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-038-073-29

```
seq_documentation_block:
; Sequence 29, Application US/09038073
; Patent No. 6194150
; GENERAL INFORMATION:
; APPLICANT: Stinchcomb, Daniel T.
; APPLICANT: Jarvis, Thale
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; INDUCTION OF GRAFT TOLERANCE
; TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
; NUMBER OF SEQUENCES: 2751
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/038,073
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,684
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-038-073-29
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-038-073-29/rev ..

Align seg 1/1 to reverse of: US-09-038-073-29 from: 1 to: 15

1 LysLysAsn 3

|||||

14 AAAAAAAT 6

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-038-073-30

seq\_documentation\_block:

; Sequence 30, Application US/09038073

; Patent No. 6194150

; GENERAL INFORMATION:

; APPLICANT: Stinchcomb, Daniel T.



Wed Jun 19 08:58:12 2002

APPLICANT: Jarvis, Thale  
APPLICANT: McSwiggen, James  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 2751  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/038.073  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/585,684  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 218/078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-038-073-30

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-038-073-30/rev ..  
Align seg 1/1 to reverse of: US-09-038-073-30 from: 1 to: 15  
1 LysLysAsn 3  
|||||||  
13 AAGAAAAAT 5







OM of: US-09-528-682-2 to: Pending\_Patents\_NA\_Main:\* out\_format : pfs

Date: Jun 18, 2002 9:42 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=rlp  
-O=/cgn2\_1/USPTO.spool/US09528682/runat\_18062002\_082444\_7874/app\_query.fasta\_1.689  
-DB=Pending\_Patents\_NA\_Main -QFMT=fastap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -GAPOP=6.000 -DEGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=0 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682\_@CGN1\_1\_8511 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT\_THREADS=1

Search information block:

Query: US-09-528-682-2

Query length: 3

Database: Pending\_Patents\_NA\_Main:\*

Database sequences: 21979536

Database length: -2067452561

Search time (sec): 7363.740000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-16352-66	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-23100-130	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-23100-145	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US084_COMB.seq	US-08-488-878A-2	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US084_COMB.seq	US-08-488-878A-3	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US088_COMB.seq	US-08-884-765A-4	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US090_COMB.seq	US-09-012-031-29	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US090_COMB.seq	US-09-012-031-29	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US090_COMB.seq	US-09-012-031B-29	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US090_COMB.seq	US-09-012-031C-29	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-834	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-1327	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-2227	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-3613	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-4055	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-5135	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-5136	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-5274	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-7037	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-7234	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-8515	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-8981	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-10245	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-10335	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-335-032-10339	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-336-376-4493	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-424-521-4	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-98	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-127	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-132	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-190	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-198	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-212	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-295	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US098B_COMB.seq	US-09-857-723-171	16.00	105.06	2.1e+04	10
/cgn2_6/ptodata/2/pna/US098B_COMB.seq	US-08-083-639C-23	16.00	104.27	2.3e+04	11
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-112	16.00	104.27	2.3e+04	11
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-167	16.00	104.27	2.3e+04	11
/cgn2_6/ptodata/2/pna/US096B_COMB.seq	US-09-628-987B-1840	16.00	104.27	2.3e+04	11

/cgn2_6/ptodata/2/pna/US097C_COMB.seq	US-09-775-479-12	16.00	104.27	2.3e+04	1
/cgn2_6/ptodata/2/pna/US097C_COMB.seq	US-09-775-479A-12	16.00	104.27	2.3e+04	1
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US00-02589-9	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US99-08537-4	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/US091_COMB.seq	US-09-151-890B-39	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/US091_COMB.seq	US-09-151-890B-39	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-238-710-9	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/US097B_COMB.seq	US-09-735-363A-83	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/US098C_COMB.seq	US-09-876-235-9	16.00	103.56	2.6e+04	12
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US98-26935-146	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US084_COMB.seq	US-08-434-705A-2	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US085_COMB.seq	US-08-510-521B-484	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US085_COMB.seq	US-08-533-582-17	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-215-436-146	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-341-399-7	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-510-378-198	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-528-209A-6205	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-528-209A-6238	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US095B_COMB.seq	US-09-531-025A-2568	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US095B_COMB.seq	US-09-531-025A-2569	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US096B_COMB.seq	US-09-636-385-2568	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US096B_COMB.seq	US-09-636-385-2569	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US096E_COMB.seq	US-09-696-347-2568	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US096E_COMB.seq	US-09-696-347-2569	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US098C_COMB.seq	US-09-877-478-6165	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US098C_COMB.seq	US-09-877-478-6166	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US099A_COMB.seq	US-09-904-567-10	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US099A_COMB.seq	US-09-904-567-11	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US099B_COMB.seq	US-09-942-310-5	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/US099B_COMB.seq	US-09-942-310-40	16.00	102.90	2.8e+04	13
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US94-07557-8	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US98-26935-89	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US98-26935-317	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US07_COMB.seq	US-07-803-627-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US07_COMB.seq	US-07-803-627-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US080_COMB.seq	US-08-087-386B-8	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US080_COMB.seq	US-08-346-456A-11	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US083_COMB.seq	US-08-359-925A-12	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US084_COMB.seq	US-08-424-921A-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US084_COMB.seq	US-08-424-921A-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US085_COMB.seq	US-08-556-355-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US085_COMB.seq	US-08-556-355-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US086_COMB.seq	US-08-656-552-14	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US091_COMB.seq	US-09-135-064-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US091_COMB.seq	US-09-135-064-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-215-436-89	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-215-436-317	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-244-889-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-244-889-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-244-889A-5	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US092_COMB.seq	US-09-244-889A-6	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-341-700A-211	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-341-700A-213	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-341-700A-255	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-341-700A-258	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-7	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-9	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-12	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-14	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-15	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-16	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-18	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US093_COMB.seq	US-09-375-673B-27	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US094_COMB.seq	US-09-475-947A-310	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-518-297-37	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-528-209A-5103	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-528-209A-5990	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/US095A_COMB.seq	US-09-528-209A-9796	16.00	102.29	3.0e+04	14
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US00-20114-25	16.00	101.73	3.2e+04	1
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-04130A-14	16.00	101.73	3.2e+04	1
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-04130A-15	16.00	101.73	3.2e+04	1
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-10742-99	16.00	101.73	3.2e+04	1
/cgn2_6/ptodata/2/pna/PCRUS_COMB.seq	PCT-US01-16352-32	16.00	101.73	3.2e+04	1















/cgn2_6/pdata/2/pna/US097B_COMB.seq:US-09-728-574-2	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3191	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US097C_COMB.seq:US-09-775-479-8	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3192	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US097A_COMB.seq:US-09-775-479-8	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3193	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098A_COMB.seq:US-09-828-534-15	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3194	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098B_COMB.seq:US-09-891-517-14	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3195	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-15	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-3196	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-16	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4047	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-17	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4048	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-18	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4049	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-19	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4050	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-891-517-20	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4051	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US099A_COMB.seq:US-09-904-744-2	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4052	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US100_COMB.seq:US-10-056-908-26	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4053	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US06019_COMB.seq:US-60-195-852-15	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4054	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US06021_COMB.seq:US-60-216-745-5592	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US096E_COMB.seq:US-09-696-791-4055	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US06021_COMB.seq:US-60-216-745-7304	16.00	100.23	3.9e+04	18	/cgn2_6/pdata/2/pna/US097B_COMB.seq:US-09-736-9559-31	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-15	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US098B_COMB.seq:US-09-830-1233-7	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-16	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099A_COMB.seq:US-09-917-138-2	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-17	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099B_COMB.seq:US-09-943-983-35	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-18	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-19	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-20	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-21	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-22	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-23	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-24	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-25	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-26	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-27	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-28	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-29	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-30	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-31	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-32	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-33	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-34	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-35	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-36	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-37	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-38	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-39	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-40	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-41	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-42	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-43	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-44	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-45	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-46	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-47	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-48	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-49	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-50	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-51	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-52	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-53	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-54	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-55	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-56	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-57	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-58	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-59	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-60	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-61	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-62	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-63	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-64	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-65	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-66	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-67	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-68	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-69	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-70	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-71	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-72	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-534-73	16.00	99.79	4.2e+04	19	/cgn2_6/pdata/2/pna/US099C_COMB.seq:US-09-996-283-33	16.00	99.79	4.2e+04	19
/cgn2_6/pdata/2/pna/US098C_COMB.seq:US-09-828-									



/cgn2_6/pdata/2/pna/US084	COMB	seq:US-08-467-939-13	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US084	COMB	seq:US-08-471-958-13	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US085	COMB	seq:US-08-509-359-177	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US085	COMB	seq:US-08-515-070-48	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US085	COMB	seq:US-08-520-946-115	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-650-093C-88	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-650-093C-89	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-650-093C-90	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-650-093C-91	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-650-093C-94	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-656-552-17	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-659-296-18	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-682-853-51	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US086	COMB	seq:US-08-683-315A-18	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US087	COMB	seq:US-08-707-395B-6	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US087	COMB	seq:US-08-784-200A-4	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US087	COMB	seq:US-08-784-200A-10	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-784-200A-19	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-784-200A-2	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-784-200A-10	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US087	COMB	seq:US-08-784-200-19	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US087	COMB	seq:US-08-784-200-19	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-870-351-6	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-870-351-16	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-870-627C-54	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-870-656-18	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-886-313-29	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-886-313-46	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US088	COMB	seq:US-08-905-772-70	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-951-831A-2	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-951-831A-6	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-951-831A-8	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-951-831A-16	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-957-709-22	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-988-321A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-988-321A-22	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-988-321A-23	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-988-321A-24	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US089	COMB	seq:US-08-989-351-11	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US090	COMB	seq:US-09-079-874-17	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US090	COMB	seq:US-09-080-140-17	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-127-238-35	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-3036	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-3055	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-4962	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-3425	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-5552	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US091	COMB	seq:US-09-198-452A-6645	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-1254	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-2841	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-2841	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-3198	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-4530	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-4607	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-4835	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-201-228A-5585	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-214-386-29	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-214-386-46	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-230-521A-74	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-230-521A-158	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-230-521-74	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-230-521-158	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-232-785-167	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-232-884-167	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-249-247-74	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US092	COMB	seq:US-09-265-503B-82	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-305-856-58	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-305-856B-58	+	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-308-845A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21	-	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US093	COMB	seq:US-09-331-568A-21					



/cgn2_6/pdata/2/pna/US095A COMB. seq: US-09-690-936-23 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US096A COMB. seq: US-09-690-936-24 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-11940 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-12010 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-12106 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-12321 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-13866 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-13866 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-16414 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-16473 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-16771 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-16897 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-16956 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-703-708-17682 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097A COMB. seq: US-09-713-681A-40 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097B COMB. seq: US-09-758-881-97 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097C COMB. seq: US-09-780-045-80 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097C COMB. seq: US-09-780-045-86 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097C COMB. seq: US-09-791-502-2 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097C COMB. seq: US-09-792-231-23 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US097C COMB. seq: US-09-794-422-36 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-802-669-62 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-808-693-11 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-8 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-9 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-10 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-11 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-12 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-13 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-14 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-15 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-16 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-17 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098A COMB. seq: US-09-823-634A-18 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-830-123-5 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-851-871-127 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-853-409-21 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-853-409-22 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-853-409-23 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098B COMB. seq: US-09-853-409-24 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-888-326-164 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-888-326-432 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-888-326-737 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-891-517-35 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-891-517-36 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-891-517-41 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-891-517-42 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-900-920-40 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US098C COMB. seq: US-09-900-920-41 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099A COMB. seq: US-09-908-158-155 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099A COMB. seq: US-09-908-159-18 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099A COMB. seq: US-09-924-125-6 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099A COMB. seq: US-09-924-439-5 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099B COMB. seq: US-09-954-679-30 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099B COMB. seq: US-09-958-236-97 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099C COMB. seq: US-09-966-284-9 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099C COMB. seq: US-09-980-052-74 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099C COMB. seq: US-09-980-052-77 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US099C COMB. seq: US-09-980-953-127 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-006-191-123 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-006-911-34 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-017-995-431 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-017-995-557 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-017-995-561 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-017-995-1091 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-031-044-29 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-056-229-321 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-088-319-22 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-09-422-978-10665 +	16.00	98.97	4.6e+04	21
/cgn2_6/pdata/2/pna/US100 COMB. seq: US-10-088-319-27 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6008 COMB. seq: US-60-082-614-2314 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6008 COMB. seq: US-60-082-614-2564 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-11940 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-12070 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-12106 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-12321 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-13786 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-13864 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-13864 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-16773 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-16897 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-16956 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6016 COMB. seq: US-60-164-320-17682 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-12016 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-12321 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-13786 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-13864 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-16414 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-16773 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-16897 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6018 COMB. seq: US-60-183-791-16956 -	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +	16.00	99.37	4.4e+04	20
/cgn2_6/pdata/2/pna/US6021 COMB. seq: US-60-216-745-5199 +				



```
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-423-058C-27 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-443-88 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-475-947A-119 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-391-15 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-391-16 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-391-17 + 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-392-15 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-392-16 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-392-17 + 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-392-17 + 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-595-770-139 - 16.00 98.97 4.6e+04 21
/cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-60-350-061-493 + 16.00 98.97 4.6e+04 21
```

seq\_name: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:PCT-US01-16352-66

## seq\_documentation\_block:

```
; Sequence 66, Application PC/TUS0116352
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc
; APPLICANT: Anastasio, Allison E
; APPLICANT: Chew, Anne
; APPLICANT: Choi, Julie Y
; APPLICANT: Denton, R. Rex
; APPLICANT: Lee, Helen H.
; APPLICANT: Nandabalan, Krishnan
; TITLE OF INVENTION: Haplotypes of the PON2 Gene
; FILE REFERENCE: MWH-0563PCT PON2
; CURRENT APPLICATION NUMBER: PCT/US01/16352
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 60/205,145
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapien
PCT-US01-16352-66
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x PCT-US01-16352-66/rev ..

Align seg 1/1 to reverse of: PCT-US01-16352-66 from: 1 to: 10

1 LysLysAsn 3

|||||  
9 AAGAAAC 1

seq\_name: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:PCT-US01-23100-130

## seq\_documentation\_block:

```
; Sequence 130, Application PC/TUS0123100
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals
; APPLICANT: Duda, Amy
; APPLICANT: Koshi, Beena
; APPLICANT: Sanchis, Angela
; TITLE OF INVENTION: Haplotypes of the CHRNA1 Gene
; FILE REFERENCE: MWH-0771PCT CHRNA1
; CURRENT APPLICATION NUMBER: PCT/US01/23100
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,538
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 10
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
PCT-US01-23100-130

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US01-23100-130 ..

Align seg 1/1 to: PCT-US01-23100-130 from: 1 to: 10

1 LysLysAsn 3
|||||
2 AAAAAAAT 10
```

seq\_name: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:PCT-US01-23100-145

```
seq_documentation_block:
; Sequence 145, Application PC/TUS0123100
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals
; APPLICANT: Duda, Amy
; APPLICANT: Koshi, Beena
; APPLICANT: Sanchis, Angela
; TITLE OF INVENTION: Haplotypes of the CHRNA1 Gene
; FILE REFERENCE: MWH-0771PCT CHRNA1
; CURRENT APPLICATION NUMBER: PCT/US01/23100
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,538
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 145
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapien
PCT-US01-23100-145
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x PCT-US01-23100-145/rev ..

Align seg 1/1 to reverse of: PCT-US01-23100-145 from: 1 to: 10

1 LysLysAsn 3

|||||  
10 AAAAAAAT 2

seq\_name: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:US-08-488-878A-2

## seq\_documentation\_block:

```
; Sequence 2, Application US/08488878A
; GENERAL INFORMATION:
; APPLICANT: Scaringe, Stephen A.
; APPLICANT: Caruthers, Marvin H.
; TITLE OF INVENTION: NOVEL PROTECTING GROUPS AND USE THEREOF
; FILE REFERENCE: MWH-0771PCT CHRNA1
; CURRENT APPLICATION NUMBER: PCT/US01/23100
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,538
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 10
; TYPE: DNA
```



```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,878A
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cleveland, Dan
; REGISTRATION NUMBER: 36,106
; REFERENCE/DOCKET NUMBER: 9028/103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)449-9497
; TELEFAX: (303)449-0814
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; US-08-488-878A-2
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
```

```
US-09-528-682-2 x US-08-488-878A-2 ..
```

```
Align seg 1/1 to: US-08-488-878A-2 from: 1 to: 10
```

```
1 LysLysAsn 3
|||||
2 AAAAAAAT 10
```

```
seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-488-878A-3
```

```
seq_documentation_block:
; Sequence 3, Application US/08488878A
; GENERAL INFORMATION:
; APPLICANT: Scarlinge, Stephen A.
; APPLICANT: Caruthers, Marvin H.
; TITLE OF INVENTION: NOVEL PROTECTING GROUPS AND USE THEREOF
; TITLE OF INVENTION: IN AN IMPROVED PROCESS FOR OLIGONUCLEOTIDE SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dan Cleveland, Jr.
; STREET: 1790 30th Street, Suite 140
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,878A
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cleveland, Dan
; REGISTRATION NUMBER: 36,106
; REFERENCE/DOCKET NUMBER: 9028/103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)449-9497
; TELEFAX: (303)449-0814
```

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (synthetic)
; US-08-488-878A-3

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-08-488-878A-3 ..

Align seg 1/1 to: US-08-488-878A-3 from: 1 to: 10

1 LysLysAsn 3
|||||
2 AAAAAAAT 10
```

```
seq_name: /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-08-864-765A-4
```

```
seq_documentation_block:
; Sequence 4, Application US/08864765A
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Peter
; APPLICANT: Knudsen, Helle
; TITLE OF INVENTION: Conjugated Peptide Nucleic Acids Having
; TITLE OF INVENTION: Enhanced Cellular Uptake
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,765A
; FILING DATE: 28-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Luccl, Joseph
; REGISTRATION NUMBER: 33,307
; REFERENCE/DOCKET NUMBER: ISIS-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: No
; FEATURE: aminoethyl glycine monomers
; US-08-864-765A-4
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```



```
alignment_block:
US-09-528-682-2 x US-08-864-765A-4/rev ..
Align seg 1/1 to reverse of: US-08-864-765A-4 from: 1 to: 10

1 LysLysAsn 3
|||||
10 AAAAAAAAT 2

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031-29

seq_documentation_block:
; Sequence 29, Application US/09012031
; GENERAL INFORMATION:
; APPLICANT: VELCULESCU, VICTOR
; APPLICANT: VOGELSTEIN, BERT
; APPLICANT: KINZLER, KENNETH
; TITLE OF INVENTION: CHARACTERIZATION OF THE YEAST
; FILE REFERENCE: 01107.73251
; CURRENT APPLICATION NUMBER: 60/035,917
; EARLIER APPLICATION NUMBER: 60/035,917
; EARLIER FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 317
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-012-031-29

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-012-031-29/rev ..
Align seg 1/1 to reverse of: US-09-012-031-29 from: 1 to: 10

1 LysLysAsn 3
|||||
9 AAGAAAAAC 1

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031B-29

seq_documentation_block:
; Sequence 29, Application US/09012031B
; GENERAL INFORMATION:
; APPLICANT: VELCULESCU, VICTOR
; APPLICANT: VOGELSTEIN, BERT
; APPLICANT: KINZLER, KENNETH
; TITLE OF INVENTION: CHARACTERIZATION OF THE YEAST
; FILE REFERENCE: 01107.73251
; CURRENT APPLICATION NUMBER: 60/035,917
; EARLIER APPLICATION NUMBER: 60/035,917
; EARLIER FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-012-031B-29

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-012-031B-29/rev ..
Align seg 1/1 to reverse of: US-09-012-031B-29 from: 1 to: 10

1 LysLysAsn 3
|||||
9 AAGAAAAAC 1

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031-29

seq_documentation_block:
; Sequence 29, Application US/09012031
; GENERAL INFORMATION:
; APPLICANT: VELCULESCU, VICTOR
; APPLICANT: VOGELSTEIN, BERT
; APPLICANT: KINZLER, KENNETH
; TITLE OF INVENTION: CHARACTERIZATION OF THE YEAST
; FILE REFERENCE: 01107.73251
; CURRENT APPLICATION NUMBER: 60/035,917
; EARLIER APPLICATION NUMBER: 60/035,917
; EARLIER FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 301
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,031
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/035,917
; FILING DATE: 23-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 01107.73251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-012-031-29

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-012-031-29/rev ..
Align seg 1/1 to reverse of: US-09-012-031-29 from: 1 to: 10

1 LysLysAsn 3
|||||
9 AAGAAAAAC 1

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031-29

seq_documentation_block:
```



seq\_name: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:US-09-012-031C-29

```
seq_documentation_block:
; Sequence 29, Application US/09012031C
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/012,031C
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: 60/035,917
; PRIOR FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-012-031C-29
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-012-031C-29/rev ..  
Align seg 1/1 to reverse of: US-09-012-031C-29 from: 1 to: 10

```
1 LysLysAsn 3
|||||
9 AAGAAAAAC 1
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-29

```
seq_documentation_block:
; Sequence 29, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-29
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-335-032-29/rev ..  
Align seg 1/1 to reverse of: US-09-335-032-29 from: 1 to: 10

```
1 LysLysAsn 3
|||||
9 AAGAAAAAC 1
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-834

```
seq_documentation_block:
; Sequence 834, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 834
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-834
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-335-032-834/rev ..  
Align seg 1/1 to reverse of: US-09-335-032-834 from: 1 to: 10

```
1 LysLysAsn 3
|||||
9 AAGAAAAAC 1
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-1327

```
seq_documentation_block:
; Sequence 1327, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1327
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-1327
```

alignment\_scores:  
Quality: 16.00 Length: 3



```
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-1327/rev ..
Align seg 1/1 to reverse of: US-09-335-032-1327 from: 1 to: 10

1 LysLysAsn 3
|||||
10 AAGAAGAAT 2

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-2227
seq_documentation_block:
; Sequence 2227, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2227
; TYPE: DNA
; LENGTH: 10
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-2227

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-2227/rev ..
Align seg 1/1 to reverse of: US-09-335-032-2227 from: 1 to: 10

1 LysLysAsn 3
|||||
10 AAGAAGAAT 2

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-3613
seq_documentation_block:
; Sequence 3613, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3613
; LENGTH: 10
```

```
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-3613

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-3613 ..
Align seg 1/1 to: US-09-335-032-3613 from: 1 to: 10

1 LysLysAsn 3
|||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-4055
seq_documentation_block:
; Sequence 4055, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4055
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-4055

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-4055/rev ..
Align seg 1/1 to reverse of: US-09-335-032-4055 from: 1 to: 10

1 LysLysAsn 3
|||||
9 AAAAAAAT 1

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-5135
seq_documentation_block:
; Sequence 5135, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
```



```
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5135
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-5135
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-335-032-5135 ..
```

```
Align seg 1/1 to: US-09-335-032-5135 from: 1 to: 10
```

```
1 LysLysAsn 3
|||||
2 AAAAAAAT 10
```

```
seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-5136
```

```
seq_documentation_block:
; Sequence 5136, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5136
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-5136
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-335-032-5136 ..
```

```
Align seg 1/1 to: US-09-335-032-5136 from: 1 to: 10
```

```
1 LysLysAsn 3
|||||
1 AAAAAAAT 9
```

```
seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-5274
```

```
seq_documentation_block:
; Sequence 5274, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
```

```
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5274
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-5274
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-335-032-5274 ..
```

```
Align seg 1/1 to: US-09-335-032-5274 from: 1 to: 10
```

```
1 LysLysAsn 3
|||||
2 AAGAAGAAT 10
```

```
seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-7067
```

```
seq_documentation_block:
; Sequence 7067, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7067
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7067
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-335-032-7067/rev ..
```

```
Align seg 1/1 to reverse of: US-09-335-032-7067 from: 1 to: 10
```

```
1 LysLysAsn 3
|||||
9 AAAAAAAT 1
```



seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-7234

```
seq_documentation_block:
; Sequence 7234, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7234
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7234
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-335-032-7234 ..  
Align seg 1/1 to: US-09-335-032-7234 from: 1 to: 10

```
1 LysLysAsn 3
|||||||
2 AAGAAAAAT 10
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-8515

```
seq_documentation_block:
; Sequence 8515, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8515
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-8515
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-335-032-8515 ..

Align seg 1/1 to: US-09-335-032-8515 from: 1 to: 10

```
1 LysLysAsn 3
|||||||
2 AAGAAAAAT 10
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-8981

```
seq_documentation_block:
; Sequence 8981, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8981
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-8981
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-335-032-8981/rev ..

Align seg 1/1 to reverse of: US-09-335-032-8981 from: 1 to: 10

```
1 LysLysAsn 3
|||||||
10 AAGAAAAAT 2
```

seq\_name: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-335-032-10245

```
seq_documentation_block:
; Sequence 10245, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10245
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-10245
```



```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-10245/rev ..
Align seg 1/1 to reverse of: US-09-335-032-10245 from: 1 to: 10

1 LysLysasn 3
|||||||
9 AAAAGAAC 1

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-10335

seq_documentation_block:
; Sequence 10335, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10335
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-10335

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-10335/rev ..
Align seg 1/1 to reverse of: US-09-335-032-10335 from: 1 to: 10

1 LysLysasn 3
|||||||
9 AAGAAAC 1

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-335-032-10339

seq_documentation_block:
; Sequence 10339, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 10339
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-10339

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-335-032-10339 ..
Align seg 1/1 to: US-09-335-032-10339 from: 1 to: 10

1 LysLysasn 3
|||||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-336-376-4493

seq_documentation_block:
; Sequence 4493, Application US/09336376
; GENERAL INFORMATION:
; APPLICANT: Roberts, Bruce L.
; TITLE OF INVENTION: POLYNUCLEOTIDE POPULATION ISOLATED FROM
; FILE REFERENCE: 126745205600
; CURRENT APPLICATION NUMBER: US/09/336.376
; CURRENT FILING DATE: 1999-06-17
; EARLIER APPLICATION NUMBER: 60/090,039
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,040
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,041
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,853
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,997
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,079
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,035
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,993
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,992
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,072
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,878
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,991
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,000
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,048
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,999
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,043
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,042
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,036
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,044
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,844
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,080
```



EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,833  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,994  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,077  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,078  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,047  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,076  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,045  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/111,715  
EARLIER FILING DATE: 1998-12-08  
NUMBER OF SEQ ID NOS: 5980  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4493  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-336-376-4493

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-336-376-4493 ..

Align seg 1/1 to: US-09-336-376-4493 from: 1 to: 10

1 LysLysAsn 3

|||||

2 AAGAAAAAC 10

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-424-521-4

## seq\_documentation\_block:

Sequence 4, Application US/09424521  
GENERAL INFORMATION:  
APPLICANT: Neilsen, Peter  
APPLICANT: Knudsen, Helle  
TITLE OF INVENTION: Conjugated Peptide Nucleic Acids Having Enhanced Cellular Uptake  
FILE REFERENCE: ISIS3070  
CURRENT APPLICATION NUMBER: US/09/424,521  
CURRENT FILING DATE: 2000-02-15  
PRIOR APPLICATION NUMBER: US 08/864,765  
PRIOR FILING DATE: 1997-05-28  
PRIOR APPLICATION NUMBER: PCT/US98/10804  
PRIOR FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Novel Sequence  
US-09-424-521-4

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-424-521-4/rev ..  
Align seg 1/1 to reverse of: US-09-424-521-4 from: 1 to: 10

1 LysLysAsn 3

|||||

10 AAAAAAAAT 2

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-98

## seq\_documentation\_block:

Sequence 98, Application US/09475947A  
GENERAL INFORMATION:  
APPLICANT: Garner, Harold R.  
APPLICANT: Wren, Jonathan D.  
APPLICANT: Minna, John D.  
TITLE OF INVENTION: Polymorphic Repeats in Human Genes  
FILE REFERENCE: UTSD0667  
CURRENT APPLICATION NUMBER: US/09/475,947A  
CURRENT FILING DATE: 1999-12-31  
NUMBER OF SEQ ID NOS: 346  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 98  
LENGTH: 10  
TYPE: DNA  
ORGANISM: human  
US-09-475-947A-98

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-475-947A-98 ..

Align seg 1/1 to: US-09-475-947A-98 from: 1 to: 10

1 LysLysAsn 3

|||||

1 AAAAAAAC 9

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-127

## seq\_documentation\_block:

Sequence 127, Application US/09475947A  
GENERAL INFORMATION:  
APPLICANT: Garner, Harold R.  
APPLICANT: Wren, Jonathan D.  
APPLICANT: Minna, John D.  
TITLE OF INVENTION: Polymorphic Repeats in Human Genes  
FILE REFERENCE: UTSD0667  
CURRENT APPLICATION NUMBER: US/09/475,947A  
CURRENT FILING DATE: 1999-12-31  
NUMBER OF SEQ ID NOS: 346  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 127  
LENGTH: 10  
TYPE: DNA  
ORGANISM: human  
US-09-475-947A-127

## alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-528-682-2 x US-09-475-947A-127 ..

Align seg 1/1 to: US-09-475-947A-127 from: 1 to: 10



```
1 LysLysAsn 3
|||||
1 AAAAAAAC 9
```

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-132

```
seq_documentation_block:
; Sequence 132, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475.947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 132
; LENGTH: 10
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-132
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:  
US-09-528-682-2 x US-09-475-947A-132 ..

Align seg 1/1 to: US-09-475-947A-132 from: 1 to: 10

```
1 LysLysAsn 3
|||||
1 AAAAAAAC 9
```

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-190

```
seq_documentation_block:
; Sequence 190, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475.947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 190
; LENGTH: 10
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-190
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:  
US-09-528-682-2 x US-09-475-947A-190 ..

Align seg 1/1 to: US-09-475-947A-190 from: 1 to: 10

```
1 LysLysAsn 3
|||||
```

```
1 AAGAAGAAC 9
```

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-198

```
seq_documentation_block:
; Sequence 198, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475.947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 198
; LENGTH: 10
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-198
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:  
US-09-528-682-2 x US-09-475-947A-198 ..

Align seg 1/1 to: US-09-475-947A-198 from: 1 to: 10

```
1 LysLysAsn 3
|||||
2 AAAAAAAC 10
```

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-212

```
seq_documentation_block:
; Sequence 212, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475.947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 212
; LENGTH: 10
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-212
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:  
US-09-528-682-2 x US-09-475-947A-212 ..

Align seg 1/1 to: US-09-475-947A-212 from: 1 to: 10

```
1 LysLysAsn 3
|||||
1 AAAAAAAC 9
```

seq\_name: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:US-09-475-947A-295



```

seq_documentation_block:
; Sequence 295, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTSD0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 295
; LENGTH: 10
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-295

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-475-947A-295 ..
Align seg 1/1 to: US-09-475-947A-295 from: 1 to: 10
1 LysLysAsn 3
|||||||
2 AAAAAAAC 10

seq_name: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-857-723-171

seq_documentation_block:
; Sequence 171, Application US/09857723
; GENERAL INFORMATION:
; APPLICANT: Alcivar-Warren, Acacia
; APPLICANT: Xu, Zhenkang
; APPLICANT: Dhar, Arun K.
; APPLICANT: Fan, Yongjun
; APPLICANT: Meehan, Dawn
; APPLICANT: Garcia, Denise K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES OF SHRIMP
; FILE REFERENCE: 1322.1023-003
; CURRENT APPLICATION NUMBER: US/09/857,723
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/111,670
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 171
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Litopenaeus vannamei
US-09-857-723-171

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-857-723-171 ..
Align seg 1/1 to: US-09-857-723-171 from: 1 to: 10
1 LysLysAsn 3
|||||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-08-083-639C-23

seq_documentation_block:
; Sequence 23, Application US/08083639C
; GENERAL INFORMATION:
; APPLICANT: Herr, John C.
; TITLE OF INVENTION: Novel Intra-Acrosomal Endoproteases
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon H. Parker, Esq.
; STREET: 300 Preston Avenue
; CITY: Charlottesville
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22902
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,639C
; FILING DATE: 28-June-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/890,433
; FILING DATE: 28-May-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Sheldon H.
; REGISTRATION NUMBER: 20,738
; REFERENCE/DOCKET NUMBER: GC-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 804-977-6606
; TELEFAX: 804-296-7605
; TELEX: Not applicable
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11bp
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: - cDNA to mRNA
; DESCRIPTION: Native SP-10 peptides isolated
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal fragment
; ORIGINAL SOURCE: A37225
; ORGANISM: Homo sapien
; STRAIN: N/A
; INDIVIDUAL ISOLATE: N/A
; DEVELOPMENTAL STAGE: Adult
; HAPLOTYPE: N/A
; TISSUE TYPE: Testis
; CELL LINE: N/A
; CELL TYPE: Protein - sperm, cDNA from
; ORGANELLE: N/A
; IMMEDIATE SOURCE:
; LIBRARY: cDNA from human testis library
; CLONE: AP-10 cDNA from human testis library
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 1
; MAP POSITION: 11 of 24
; UNITS: N/A
; FEATURE:
; NAME/KEY: SP-10 is a unique sperm protein found
; LOCATION: N/A
; IDENTIFICATION METHOD: SP-10 is unique
; OTHER INFORMATION: Product appears to be
; PUBLICATION INFORMATION:
; AUTHORS: Herr, John C.
; TITLE: Purification and Microsequencing of the
; JOURNAL: Biology of Reproduction

```



```
; VOLUME: 47
; ISSUE:
; PAGES: 11-20
; DATE: 1992
; US-08-083-639C-23

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-08-083-639C-23/rev ..

Align seg 1/1 to reverse of: US-08-083-639C-23 from: 1 to: 11

1 LysLysAsn 3
|||||
11 AAAAAAAT 3

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-475-947A-112

seq_documentation_block:
; Sequence 112, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 112
; LENGTH: 11
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-112

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-09-475-947A-112 ..

Align seg 1/1 to: US-09-475-947A-112 from: 1 to: 11

1 LysLysAsn 3
|||||
2 AAAAAAAT 10

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-475-947A-167

seq_documentation_block:
; Sequence 167, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 167
; LENGTH: 11
; TYPE: DNA
```

```
; ORGANISM: human
US-09-475-947A-167

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-09-475-947A-167 ..

Align seg 1/1 to: US-09-475-947A-167 from: 1 to: 11

1 LysLysAsn 3
|||||
1 AAAAAAAT 9

seq_name: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-628-987B-1840

seq_documentation_block:
; Sequence 1840, Application US/09628987B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1075P
; CURRENT APPLICATION NUMBER: US/09/628,987B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 1862
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1840
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..11
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: 1..11
; OTHER INFORMATION: Ceres Seq. ID 1447829
US-09-628-987B-1840

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-2 x US-09-628-987B-1840 ..

Align seg 1/1 to: US-09-628-987B-1840 from: 1 to: 11

1 LysLysAsn 3
|||||
2 AAGAAGAAC 10

seq_name: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-775-479-12

seq_documentation_block:
; Sequence 12, Application US/09775479
; GENERAL INFORMATION:
; APPLICANT: LECLERC, Guy
; APPLICANT: MARTEL, R.mi
; TITLE OF INVENTION: RADIOLABELED DNA CARRIER, METHOD OF
; TITLE OF INVENTION: RADIOLABELED DNA CARRIER, METHOD OF PREPARATION AND
; FILE REFERENCE: 12168-1US-2
; CURRENT APPLICATION NUMBER: US/09/775,479
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/318,106
```



```
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: 08/756,728
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide
US-09-775-479-12

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-775-479-12 ..
Align seg 1/1 to: US-09-775-479-12 from: 1 to: 11

1 LysLysAsn 3
1 AAAAAAAAAAT 9

seq_name: /cgn2_5/ptodata/2/pna/US097C_COMB.seq:US-09-775-479A-12

seq_documentation_block:
; Sequence 12, Application US/09775479A
; GENERAL INFORMATION:
; APPLICANT: LECLERC, Guy
; TITLE OF INVENTION: RADIO-LABELED DNA CARRIER, METHOD OF
; TITLE OF INVENTION: RADIO-LABELED DNA CARRIER, METHOD OF PREPARATION AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; FILE REFERENCE: 01826-50018 CIP
; CURRENT APPLICATION NUMBER: US/09/775,479A
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/775,479
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/318,106
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: 08/756,728
; PRIOR FILING DATE: 1996-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide
US-09-775-479A-12

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-775-479A-12 ..
Align seg 1/1 to: US-09-775-479A-12 from: 1 to: 11

1 LysLysAsn 3
1 AAAAAAAAAAT 9

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US00-02589-9

seq_documentation_block:
; Sequence 4, Application PC/TUS9906537A
; GENERAL INFORMATION:
; APPLICANT: Feng, Feng
; TITLE OF INVENTION: Identifying Ligands of Target Proteins With Target
; TITLE OF INVENTION: Complementary Library Technology (TCLT)
; FILE REFERENCE: 019815-000200PC
; CURRENT APPLICATION NUMBER: PCT/US99/06537A
; CURRENT FILING DATE: 1999-04-19
; EARLIER APPLICATION NUMBER: US 60/083,046
; EARLIER FILING DATE: 1998-04-24
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:anti-sense
PCT-US99-06537-4

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US99-06537-4 ..
Align seg 1/1 to: PCT-US99-06537-4 from: 1 to: 12

1 LysLysAsn 3
```



```
|||||
1 AAAAAAGAC 9

seq_name: /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-151-890-39

seq_documentation_block:
; Sequence 39, Application US/09151890
; GENERAL INFORMATION:
; APPLICANT: Gary P. Schroth
; APPLICANT: Thomas Wayne Bruce
; APPLICANT: Young J. Suh
; FILE OF INVENTION: Nucleic Acid Ligand Interaction Assays
; FILE REFERENCE: 4600-0128
; CURRENT APPLICATION NUMBER: US/09/151.890
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
US-09-151-890-39

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-151-890-39 ..
Align seg 1/1 to: US-09-151-890-39 from: 1 to: 12

1 LysLysAsn 3
|||||
3 AAAAAAAGAC 11

seq_name: /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-151-890B-39

seq_documentation_block:
; Sequence 39, Application US/09151890B
; GENERAL INFORMATION:
; APPLICANT: Gary P. Schroth
; APPLICANT: Thomas Wayne Bruce
; APPLICANT: Young J. Suh
; FILE OF INVENTION: Nucleic Acid Ligand Interaction Assays
; FILE REFERENCE: 4600-0128
; CURRENT APPLICATION NUMBER: US/09/151.890B
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
US-09-151-890B-39

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-151-890B-39 ..
Align seg 1/1 to: US-09-151-890B-39 from: 1 to: 12

1 LysLysAsn 3
|||||
3 AAAAAAAGAC 11

seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-735-363A-83

seq_documentation_block:
; Sequence 83, Application US/09735363A
; GENERAL INFORMATION:
; APPLICANT: Fillon, Mario
; APPLICANT: Phillip, Nigel
; FILE OF INVENTION: Therapeutically Useful Synthetic Oligonucleotides
; FILE REFERENCE: 02811-0181
; CURRENT APPLICATION NUMBER: US/09/735,363A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/170,325
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 60/228,925
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 83
; LENGTH: 12
```



```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-735-363A-83

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-735-363A-83/rev ..
Align seg 1/1 to reverse of: US-09-735-363A-83 from: 1 to: 12

1 LysLysAsn 3
|||||
11 AAAAAAAC 3

seq_name: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-876-235-9

seq_documentation_block:
; Sequence 9, Application US/09876235
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350005
; CURRENT APPLICATION NUMBER: US/09/876,235
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/247,190
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,963
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/064,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/007,005
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-14
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 9
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
US-09-876-235-9

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-876-235-9 ..
Align seg 1/1 to: US-09-876-235-9 from: 1 to: 12

1 LysLysAsn 3
|||||
3 AAAAAAAC 11

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US98-26935-146

seq_documentation_block:
; Sequence 146, Application PC/TUS9826935
; GENERAL INFORMATION:
; APPLICANT: Jessen, Holly
; APPLICANT: Webb, David
; APPLICANT: Corvett, Virginia H.
; APPLICANT: Schupp, James M.
; APPLICANT: Keim, Paul S.
; APPLICANT: Pioneer Hi-Bred International, Inc.
; TITLE OF INVENTION: Nucleotide Polymorphisms in Soybean
; FILE REFERENCE: 018574-000110PC
; CURRENT APPLICATION NUMBER: PCT/US98/26935
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: US 60/068,185
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 146
; LENGTH: 13
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: PCT-US98-26935-146

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US98-26935-146 ..
Align seg 1/1 to: PCT-US98-26935-146 from: 1 to: 13

1 LysLysAsn 3
|||||
1 AAAAAAAC 9
```











OM of: US-09-528-682-2 to: Pending\_Patents\_NA\_New:\* out\_format : pfs

Date: Jun 18, 2002 9:49 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=cg2\_1/USPTO.spool/US09528682/runat\_18062002\_082444\_7900/app\_query.fasta\_1.689  
-DB=pending\_Patents\_NA\_New -QFWT=fastap -SURFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALLIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682.cg2n1\_1572 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-2

Query length: 3

Database: Pending\_Patents\_NA\_New:\*

Database sequences: 1014543

Database length: 727792371

Search time (sec): 403.930000

score\_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: PCT-US02-15111-28 +	16.00	101.27	1.6e+03	12	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: PCT-US02-15111-27 +	16.00	100.66	1.7e+03	13	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: PCT-US02-15111-30 +	16.00	100.66	1.7e+03	13	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-146 +	16.00	100.66	1.7e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-89 +	16.00	100.10	1.8e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-317 +	16.00	100.10	1.8e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-875-433A-103 -	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-889-866A-54 -	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-889-866A-58 -	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-889-866A-60 -	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-76 +	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-346 +	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-383 +	16.00	99.58	2.0e+03	1	
/cg2_6/ptodata/1/pna/US08 NEW COMB. seq: US-10-112-653-833 -	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-131-102-7 -	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-131-102-8 -	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-131-102-9 -	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-861-787A-1 +	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-861-787A-2 +	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-861-787A-3 +	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-861-787A-4 +	16.00	99.09	2.1e+03	16	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-581-970A-69 -	16.00	99.09	2.1e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-941-492-16 -	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-58 -	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-59 +	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-60 +	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-63 +	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-240 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-241 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-242 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-243 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-244 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-245 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-917 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-918 -	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-745-237A-1448 +	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-043-875-54 +	16.00	98.63	2.2e+03	17	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-138-674-1319 +	16.00	98.63	2.2e+03	1	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-138-674-6287 +	16.00	98.63	2.2e+03	1	

/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-138-674-8182 +	16.00	98.63	2.2e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-138-674-8315 +	16.00	98.63	2.2e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-10-138-674-8601 +	16.00	98.63	2.2e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-14 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-15 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-16 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-17 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-18 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-19 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-20 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-61 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-62 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-959-964-222 -	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US08 NEW COMB. seq: US-10-043-875-55 +	16.00	98.20	2.3e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-08-803-331-29 +	16.00	97.79	2.5e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-943-983C-35 +	16.00	97.79	2.5e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-923-327-70 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-632-702-22 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-24 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-25 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-29 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-30 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-31 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-35 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-725-265-36 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-3 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-3 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-31 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-31 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-44 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-442-054A-44 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-631-613-22 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-8 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-9 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-10 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-11 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-12 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-13 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-14 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-15 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-16 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-17 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-823-647B-18 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-882-945A-257 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-932-257A-29 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-932-257A-29 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-112-653-413 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-112-653-538 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-112-653-1035 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-121-120-157 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-023-58334 +	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-033-297-40 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-10-130-080-27 -	16.00	97.40	2.6e+03	
/cg2_6/ptodata/1/pna/US06 NEW COMB. seq: US-60-368-184-21 -	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-551 +	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-751 -	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-764 +	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-785 +	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-786 +	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/US10 NEW COMB. seq: US-09-940-244-406 -	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: US-10-116-949-88 -	16.00	97.04	2.7e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US02-09559-2 +	16.00	96.68	2.9e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-50 +	16.00	96.68	2.9e+03	
/cg2_6/ptodata/1/pna/PCT NEW COMB. seq: PCT-US01-44838-938 +	16.00	96.68	2.9e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-909-567B-27 -	16.00	96.68	2.9e+03	
/cg2_6/ptodata/1/pna/US09 NEW COMB. seq: US-09-976-618A-45 +	16.00	96.68	2.9e+03	



/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-618A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54720	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-900A-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54795	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-900A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54931	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-974-500A-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54932	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-974-500A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54933	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-863A-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54934	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-863A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54935	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-693-005A-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54936	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-693-005A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54937	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-966-312-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54938	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-966-312-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54939	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-875-453A-84	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54940	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-875-453A-92	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54941	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-442-054A-85	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54942	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-376A-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54943	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-376A-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54944	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-66A	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54945	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-749-2	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54946	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-107-871-2	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54947	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-132-561-9	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54948	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-005-469-11	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54949	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-008-978-45	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54950	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-008-978-47	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54951	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-008-978-74	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54952	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-008-978-75	+ 16.00	96.68	2.9e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54953	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-4438B-1260	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54954	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-698-495-6	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54955	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-685-209A-6	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54956	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-685-404A-6	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54957	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-291-417B-64	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54958	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-685-913A-6	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54959	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-029-397A-7	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54960	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-112	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54961	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-234	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54962	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-523	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54963	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-539	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54964	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-913	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54965	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-029-397A-7	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54966	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-916	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54967	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-919	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54968	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-518B2	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54969	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-069-141-6	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54970	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-069-141-10	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54971	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-069-141-18	+ 16.00	96.35	3.0e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54972	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-442-489B-55	+ 16.00	96.03	3.1e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54973	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-653-913	+ 16.00	96.03	3.1e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54974	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-116-288-15	+ 16.00	96.03	3.1e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54975	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-07487-54	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54976	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-07487-55	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54977	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-101	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54978	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-5412	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54979	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-5958	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54980	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-5962	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54981	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-5969	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-54982	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77337	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77338	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-8144	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77339	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-80959	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77340	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-20959	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77341	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-22301	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77342	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-22969	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77343	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-25954	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77344	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-25803	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77345	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26971	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77346	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26986	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77347	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26989	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77348	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26995	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77349	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26997	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77350	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26998	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77351	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-26999	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77352	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-27000	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77353	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-27001	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77354	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-27002	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77355	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-27003	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-77356	+ 16.00	95.72	3.2e+03	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-956-604-27004	+ 16.00	95.72	3.2e+03	2	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-				



[illegible]



```
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-066-409-9 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-066-409-10 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-066-409-11 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-066-409-12 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-176812 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-004-427-6 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-004-427-10 - 16.00 92.16 5.1e+03 40
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-004-427-13 - 16.00 91.80 5.3e+03 42
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-150-407-3 - 16.00 91.80 5.3e+03 42
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-031-241-27 - 16.00 91.80 5.3e+03 42
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-15111-135 - 16.00 91.62 5.5e+03 43
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-940-244-97 + 16.00 91.62 5.5e+03 43
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-033-297-97 + 16.00 91.62 5.5e+03 43
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-503-138B-3 - 16.00 91.44 5.6e+03 44
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-882-945A-41 - 16.00 91.44 5.6e+03 44
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-503-138B-105 + 16.00 91.44 5.6e+03 44
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-103-613-192 - 16.00 91.44 5.6e+03 44
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-150-407-102 + 16.00 91.27 5.7e+03 45
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13873-32 - 16.00 91.27 5.7e+03 45
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13873-33 + 16.00 91.27 5.7e+03 45
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-578-519B-52 - 16.00 91.27 5.7e+03 46
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-013-803A-10 - 16.00 91.27 5.7e+03 46
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08850-2 - 16.00 91.11 5.8e+03 46
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12612-42 - 16.00 91.11 5.8e+03 46
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-127-890-21 - 16.00 91.11 5.8e+03 46
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-125-968-42 - 16.00 91.11 5.8e+03 46
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-055-728-32 - 16.00 91.11 5.8e+03 46
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-52411 + 16.00 91.11 5.8e+03 47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-52426 + 16.00 91.11 5.8e+03 47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-160-237-5 - 16.00 91.11 5.8e+03 47
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08850-3 - 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-54 + 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-71 - 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-72 + 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-103-119 - 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-104-54 + 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-704-71 - 16.00 90.94 6.0e+03 47
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08850-4 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12612-1409 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-775-217A-4 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-344-783C-27 + 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-823-647B-7 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-33 + 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-33 + 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-125-968-1409 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-704-33 + 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-176009 - 16.00 90.79 6.1e+03 48
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08850-5 - 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08850-6 - 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-720-979-2 + 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-720-979A-2 + 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-244A-9 - 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-244A-45 - 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-087-286-31 - 16.00 90.63 6.2e+03 49
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-997-672-29 - 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-196-447B-11 - 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-780-863-59 + 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-724-857-25 - 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-624-36 - 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-177831 + 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-082-894-7 - 16.00 90.48 6.3e+03 50
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-445-614A-3 - 16.00 90.33 6.4e+03 51
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-823-647B-2 - 16.00 90.33 6.4e+03 51
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-686-838B-32 + 16.00 90.33 6.4e+03 51
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-133-545-2 - 16.00 90.33 6.4e+03 51
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-133-036-16 + 16.00 90.33 6.4e+03 51
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-2767 - 16.00 90.18 6.6e+03 52
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-4132 + 16.00 90.18 6.6e+03 52
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-041-007-31 - 16.00 90.18 6.6e+03 52
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-633-75798 - 16.00 90.18 6.6e+03 52
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12612-112 + 16.00 90.04 6.7e+03 53
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-982-667-63 - 16.00 90.04 6.7e+03 53
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-940-244-52 + 16.00 90.04 6.7e+03 53
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-041-007-40 - 16.00 90.04 6.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-125-968-112 + 16.00 90.04 6.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-033-297-52 + 16.00 90.04 6.7e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-900-345A-164 - 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-2823 - 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-113-573-1 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-53525 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-10901 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-10919 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-13404 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-13476 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-13477 + 16.00 89.89 6.8e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-949-134A-8 + 16.00 89.76 6.9e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-47593-18 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-08908-9 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-10421-366 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-344-783C-30 - 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-192-9 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-3028 - 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-3073 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-699-366 + 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-041-007-11 - 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-041-007-12 - 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-178576 - 16.00 89.62 7.1e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-773-861B-1 + 16.00 89.49 7.2e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-823-647B-5 - 16.00 89.49 7.2e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-089-595-6 - 16.00 89.49 7.2e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12612-1410 - 16.00 89.35 7.3e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-097-105-838 + 16.00 89.35 7.3e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-125-968-2921 - 16.00 89.35 7.3e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-1410 - 16.00 89.35 7.3e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-189-948 - 16.00 89.23 7.4e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-127-391-26 - 16.00 89.23 7.4e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-005-338A-58 + 16.00 89.23 7.4e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12378-329 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-12378-530 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-12543 - 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-19556 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-20841 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-29828 - 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-09-651-685A-10 - 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-2852 - 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-113-872-1988 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-113-872-1989 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-113-872-1990 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-124-805-529 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-124-805-530 + 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-146-502-335 - 16.00 89.10 7.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-695-437A-49 - 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-695-437A-51 - 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-20724 + 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-2976 - 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-3051 - 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-3114 - 16.00 88.97 7.7e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-57831 + 16.00 88.85 7.8e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-23700 + 16.00 88.85 7.8e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-26823 + 16.00 88.85 7.8e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-35873 - 16.00 88.85 7.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-3148 + 16.00 88.85 7.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-178504 - 16.00 88.73 7.9e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-27908 - 16.00 88.73 7.9e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-09-540-210B-32370 + 16.00 88.73 7.9e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-7662 + 16.00 88.73 7.9e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-11453 + 16.00 88.73 7.9e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-613-169 + 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-14372-24 - 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-14081 + 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-3890 + 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-5535 - 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-1951 - 16.00 88.61 8.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-806B-2028 + 16.00 88.61 8.0e+03
```







[illegible]







```
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-178039 - 16.00 84.52 1.4e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-178053 - 16.00 84.52 1.4e+04
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-377-240-10850 - 16.00 84.52 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-2108-32570 + 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-2768 + 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-13287 - 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-689-909-737 + 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-732-560-42 - 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-16265 + 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-460-592B-938 + 16.00 84.45 1.4e+04
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-377-240-1508 + 16.00 84.45 1.4e+04

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-15111-28

seq_documentation_block:
; Sequence 28, Application PC/TUS0215111
; GENERAL INFORMATION:
; APPLICANT: Replicdyne, Inc.
; TITLE OF INVENTION: System for Discovery of Agents that Block Yersinia Pestis and Pse
; FILE REFERENCE: RDYN.02
; CURRENT APPLICATION NUMBER: PCT/US02/15111
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 60/290,725
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/332,644
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 28
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mutation
PCT-US02-15111-28

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US02-15111-28 ..
Align seg 1/1 to: PCT-US02-15111-28 from: 1 to: 12

1 LysLysAsn 3
2 AAAAAGAAAT 10

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-15111-27

seq_documentation_block:
; Sequence 27, Application PC/TUS0215111
; GENERAL INFORMATION:
; APPLICANT: Replicdyne, Inc.
; TITLE OF INVENTION: System for Discovery of Agents that Block Yersinia Pestis and Pse
; FILE REFERENCE: RDYN.02
; CURRENT APPLICATION NUMBER: PCT/US02/15111
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 60/290,725
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/332,644
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 13
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: mutation
PCT-US02-15111-27

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US02-15111-27 ..
Align seg 1/1 to: PCT-US02-15111-27 from: 1 to: 13

1 LysLysAsn 3
2 AAAAAGAAAT 10

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-15111-30

seq_documentation_block:
; Sequence 30, Application PC/TUS0215111
; GENERAL INFORMATION:
; APPLICANT: Replicdyne, Inc.
; TITLE OF INVENTION: System for Discovery of Agents that Block Yersinia Pestis and
; FILE REFERENCE: RDYN.02
; CURRENT APPLICATION NUMBER: PCT/US02/15111
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 60/290,725
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/332,644
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 13
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mutation
PCT-US02-15111-30

alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x PCT-US02-15111-30 ..
Align seg 1/1 to: PCT-US02-15111-30 from: 1 to: 13

1 LysLysAsn 3
2 AAGAAAAAAC 10

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-146

seq_documentation_block:
; Sequence 146, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORVELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
```



```

: PRIOR APPLICATION NUMBER: PCT/US98/26935
: PRIOR FILING DATE: 1998-12-18
: PRIOR APPLICATION NUMBER: 60/068,185
: PRIOR FILING DATE: 1997-12-19
: NUMBER OF SEQ ID NOS: 798
: SOFTWARE: PatentIn Ver 2.1
: SEQ ID NO 146
: LENGTH: 13
: TYPE: DNA
: ORGANISM: Glycine max
: US-09-581-970A-146

```

```
alignment_scores:      Length: 16.00      3
                       Ratio:  5.333      0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-581-970A-146
..
```

Align seg 1/1 to: US-09-581-970A-146 from: 1 to: 13

```

1 LysLysAsn 3
  |||||
1 AAAAAAAT 9

```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-581-970A-89

```

seq_documentation_block:
; Sequence 89, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORVELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHIS
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 89
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Glycine max
US-09-581-970A-89

```

alignment_scores:	
Quality:	16.00
Ratio:	5.333
Percent Similarity:	100.000
Percent Identity:	100.000
Gaps:	0
Length:	3

```
alignment_block:      x US-09-581-970A-89/rev
US-09-528-682-2
```

Align seq 1/1 to reverse of: US-09-581-970A-89 from: 1 to: 14

```

1  LysLysAsn 3
   |||||
13 AAAAAAAT 5

```

seq\_name: /cqn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-581-970A-317

```
seq_documentation_block:
; Sequence 317, Application US/09581970A
; GENERAL INFORMATION:
```

```

; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORYELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KETIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 317
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Glycine max
US-09-581-970A-317

```

```
alignment_scores:      Quality: 16.00      Length: 3
                       Ratio: 5.333      Gaps: 0
percent Similarity: 100.000
percent Identity: 100.000
```

alignment\_block:  
US-09-528-682-2 x US-09-581-970A-317

Align seq 1/1 to: US-09-581-970A-317 from: 1 to: 14

1 LysLysAsn 3  
|||  
3 AAGAAAAAT 11

seq\_name: /cqn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-875-453A-103

```

seq_documentation_block:
; Sequence 103, Application US/09875453A
; GENERAL INFORMATION:
; APPLICANT: Kim, Jungsuuh P.
; APPLICANT: Starr, Douglas B.
; APPLICANT: Tam, Albert W.
; APPLICANT: Laurance, Megan E.
; APPLICANT: Michelotti, Emil F.
; APPLICANT: Velligan, Mark D.
; APPLICANT: Latour, Derek R.
; APPLICANT: Thomas, Rita L.
; APPLICANT: Kongpachith, Ana
; APPLICANT: Sheppard, Liana T.
; APPLICANT: Lim, Moon Young
; APPLICANT: Bruce, Thomas W.
; TITLE OF INVENTION: PROMOTERS FOR REGULATED GENE EXPRESSION
; FILE REFERENCE: 54600-8135-US00
; CURRENT APPLICATION NUMBER: US/09/875,453A
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 05 60/209,549
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mutant sequence
US-09-875-453A-103

```

alignment_scores:	Quality:	Length:
	16.00	Gaps:
	Ratio: 5.333	3
		0



```
Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-875-453A-103/rev  ..
Align seg 1/1 to reverse of: US-09-875-453A-103 from: 1 to: 14

1 LysLysAsn 3
11 AAAAAAAC 3

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-889-866A-54
seq_documentation_block:
; Sequence 54, Application US/09889866A
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Denton, R. Rex
; APPLICANT: Kliem, Stefanie E.
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
; TITLE OF INVENTION: IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE
; FILE REFERENCE: MWH-0008US IGERB
; CURRENT APPLICATION NUMBER: US/09/889,866A
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/150,423
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/22175
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 145
; SEQ ID NO 54
; SOFTWARE: PatentIn Ver. 2.1 PLUS editing text file in Notepad
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-889-866A-54

alignment_scores:
Quality: 16.00    Length: 3
Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-889-866A-54/rev  ..
Align seg 1/1 to reverse of: US-09-889-866A-54 from: 1 to: 15

1 LysLysAsn 3
14 AAAAAAAAT 6

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-889-866A-58
seq_documentation_block:
; Sequence 58, Application US/09889866A
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Denton, R. Rex
; APPLICANT: Kliem, Stefanie E.
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
; TITLE OF INVENTION: IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE
; FILE REFERENCE: MWH-0008US IGERB
; CURRENT APPLICATION NUMBER: US/09/889,866A
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/150,423
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/22175
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 145
```

```
; SOFTWARE: PatentIn Ver. 2.1 PLUS editing text file in Notepad
; SEQ ID NO 58
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-889-866A-58

alignment_scores:
Quality: 16.00    Length: 3
Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-889-866A-58/rev  ..
Align seg 1/1 to reverse of: US-09-889-866A-58 from: 1 to: 15

1 LysLysAsn 3
10 AAAAAAAAT 2

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-889-866A-60
seq_documentation_block:
; Sequence 60, Application US/09889866A
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Denton, R. Rex
; APPLICANT: Kliem, Stefanie E.
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
; TITLE OF INVENTION: IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE
; FILE REFERENCE: MWH-0008US IGERB
; CURRENT APPLICATION NUMBER: US/09/889,866A
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/150,423
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/22175
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn Ver. 2.1 PLUS editing text file in Notepad
; SEQ ID NO 60
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-889-866A-60

alignment_scores:
Quality: 16.00    Length: 3
Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-889-866A-60/rev  ..
Align seg 1/1 to reverse of: US-09-889-866A-60 from: 1 to: 15

1 LysLysAsn 3
10 AAAAAAAAT 2

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-76
seq_documentation_block:
; Sequence 76, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORYELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
```



```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-346
seq_documentation_block:
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 76
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Glycine max
US-09-581-970A-76
..
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-581-970A-76
..
Align seg 1/1 to: US-09-581-970A-76 from: 1 to: 15
  1 LysLysAsn 3
  |||||
  1 AAAAAAAC 9
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-346
seq_documentation_block:
; Sequence 346, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORVELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 346
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Glycine max
US-09-581-970A-346
..
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-581-970A-346
..
Align seg 1/1 to: US-09-581-970A-346 from: 1 to: 15
  1 LysLysAsn 3
  |||||
  1 AAAAAAAC 9
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-383
seq_documentation_block:
; Sequence 383, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORVELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO 383
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Glycine max
US-09-581-970A-383
..
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-09-581-970A-383
..
Align seg 1/1 to: US-09-581-970A-383 from: 1 to: 15
  1 LysLysAsn 3
  |||||
  4 AAAAAAAT 12
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-653-833
seq_documentation_block:
; Sequence 833, Application US/10112653
; GENERAL INFORMATION:
; APPLICANT: Arleg, Arthur M.
; APPLICANT: Berg, Daniel J.
; TITLE OF INVENTION: IMMUNOSTIMULATORY NUCLEIC ACID FOR
; FILE REFERENCE: C01039/70060(AWS)
; CURRENT APPLICATION NUMBER: US/10/112,653
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/279,642
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 1040
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 833
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-112-653-833
..
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-2 x US-10-112-653-833/rev
..
```



Align seq 1/1 to reverse of: US-10-112-653-833 from: 1 to: 15

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:US-08-802-331-28

seq\_documentation\_block:  
; Sequence 28, Application US/08802331  
; GENERAL INFORMATION:  
; APPLICANT: Cook, Phillip D.  
; APPLICANT: Monia, Brett  
; APPLICANT: Martin, Pierre  
; APPLICANT: Altman, Karl-Heinz  
; TITLE OF INVENTION: Sugar-Modified Gapped Oligonucleotides  
; FILE REFERENCE: ISN00083  
; CURRENT APPLICATION NUMBER: US/08/802,331  
; CURRENT FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Novel Sequence  
US-08-802-331-28

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-08-802-331-28/rev ..

Align seq 1/1 to reverse of: US-08-802-331-28 from: 1 to: 16

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-131-102-7

seq\_documentation\_block:  
; Sequence 7, Application US/09131102  
; GENERAL INFORMATION:  
; APPLICANT: Cook, Phillip D.  
; APPLICANT: Manoharan, Muthiah  
; APPLICANT: Bhat, Balkrishen  
; TITLE OF INVENTION: Oligonucleotide Analogs Having Modified Dimers  
; FILE REFERENCE: ISIS-2906  
; CURRENT APPLICATION NUMBER: US/09/131,102  
; CURRENT FILING DATE: 1998-08-07  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (5)..(5)  
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety

; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (6)..(6)  
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (8)..(8)  
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (9)..(9)  
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (10)..(10)  
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (12)..(12)  
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (13)..(13)  
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety  
US-09-131-102-7

alignment\_scores:

Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-131-102-7/rev ..

Align seq 1/1 to reverse of: US-09-131-102-7 from: 1 to: 16

1 LysLysAsn 3  
|||||  
11 AAAAAAAC 3

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-131-102-8

seq\_documentation\_block:  
; Sequence 8, Application US/09131102  
; GENERAL INFORMATION:  
; APPLICANT: Cook, Phillip D.  
; APPLICANT: Manoharan, Muthiah  
; APPLICANT: Bhat, Balkrishen  
; TITLE OF INVENTION: Oligonucleotide Analogs Having Modified Dimers  
; FILE REFERENCE: ISIS-2906  
; CURRENT APPLICATION NUMBER: US/09/131,102  
; CURRENT FILING DATE: 1998-08-07  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Methylene (methylimino) backbone, 2'-O-CH3 ribose sugar moiety



```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: Methylene (methylimino) backbone, 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Methylene (methylimino) backbone, 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Methylene (methylimino) backbone, 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: 2'-O-CH3 Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: 2'-O-CH3 ribose sugar moiety
US-09-131-102-8

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-131-102-8 rev ..
Align seg 1/1 to reverse of: US-09-131-102-8 from: 1 to: 16

  1 LysLysasn 3
  |||||
  11 AAAAAAAC 3

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-131-102-9

seq_documentation_block:
; Sequence 9, Application US/09131102
; GENERAL INFORMATION:
; APPLICANT: Cook, Phillip D.
; APPLICANT: Manoharan, Muthiah
; APPLICANT: Bhat, Balkrishen
; TITLE OF INVENTION: Oligonucleotide Analogs Having Modified Dimers
; FILE REFERENCE: ISIS-2906
; CURRENT APPLICATION NUMBER: US/09/131.102
; CURRENT FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Methylene (methylimino) backbone, 2'-O-CH3 ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: Ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: Methylene (methylimino) backbone, ribose sugar moiety
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: Ribose sugar moiety
US-09-131-102-9

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-131-102-9 rev ..
Align seg 1/1 to reverse of: US-09-131-102-9 from: 1 to: 16

  1 LysLysasn 3
  |||||
  11 AAAAAAAC 3

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-861-787A-1

seq_documentation_block:
; Sequence 1, Application US/09861787A
; GENERAL INFORMATION:
; APPLICANT: Virtanen, Jorma
; TITLE OF INVENTION: Gene Sequencer and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
```



```

;
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows 98
; SOFTWARE: MS Word
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861,787A
; FILING DATE: 21-May-2001
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/564,399
; FILING DATE: May 1, 2000
; APPLICATION NUMBER: 09/064,635
; FILING DATE: April 21, 1998
; APPLICATION NUMBER: PCT/US/98/03362
; FILING DATE: February 20, 1998
; APPLICATION NUMBER: 60/039,027
; FILING DATE: February 21, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITH, Guy P.
; REGISTRATION NUMBER: 20,142
; REFERENCE/DOCKET NUMBER: 18950-67 (formerly 18950-62)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 788-5100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded DNA
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-861-787A-1

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-861-787A-1 ..

Align seg 1/1 to: US-09-861-787A-1 from: 1 to: 16

1 LysLysAsn 3
|||||
5 AAAAAAAC 13

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-861-787A-2

seq_documentation_block:
; Sequence 2, Application US/09861787A
; GENERAL INFORMATION:
; APPLICANT: Virtanen, Jorma
; TITLE OF INVENTION: Gene Sequencer and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows 98
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861,787A
; FILING DATE: 21-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/564,399

```

```

;
; FILING DATE: May 1, 2000
; APPLICATION NUMBER: 09/064,635
; FILING DATE: April 21, 1998
; APPLICATION NUMBER: PCT/US/98/03362
; FILING DATE: February 20, 1998
; APPLICATION NUMBER: 60/039,027
; FILING DATE: February 21, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITH, Guy P.
; REGISTRATION NUMBER: 20,142
; REFERENCE/DOCKET NUMBER: 18950-67 (formerly 18950-62)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 788-5100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded DNA
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-861-787A-2

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-861-787A-2 ..

Align seg 1/1 to: US-09-861-787A-2 from: 1 to: 16

1 LysLysAsn 3
|||||
5 AAAAAAAC 13

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-861-787A-3

seq_documentation_block:
; Sequence 3, Application US/09861787A
; GENERAL INFORMATION:
; APPLICANT: Virtanen, Jorma
; TITLE OF INVENTION: Gene Sequencer and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows 98
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861,787A
; FILING DATE: 21-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/564,399
; FILING DATE: May 1, 2000
; APPLICATION NUMBER: 09/064,635
; FILING DATE: April 21, 1998
; APPLICATION NUMBER: PCT/US/98/03362
; FILING DATE: February 20, 1998
; APPLICATION NUMBER: 60/039,027
; FILING DATE: February 21, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITH, Guy P.

```



```
;
; REGISTRATION NUMBER: 20,142
; REFERENCE/DOCKET NUMBER: 18950-67 (formerly 18950-62)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 788-5100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded DNA
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-861-787A-3

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-861-787A-3 ..
Align seg 1/1 to: US-09-861-787A-3 from: 1 to: 16

1 LysLysAsn 3
|||||
5 AAAAAAAT 13

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-861-787A-4

seq_documentation_block:
; Sequence 4, Application US/09861787A
; GENERAL INFORMATION:
; APPLICANT: Virtanen, Jorma
; TITLE OF INVENTION: Gene Sequencer and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows 98
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861,787A
; FILING DATE: 21-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/564,399
; FILING DATE: May 1, 2000
; APPLICATION NUMBER: 09/064,635
; FILING DATE: April 21, 1998
; APPLICATION NUMBER: PCT/US/98/03362
; FILING DATE: February 20, 1998
; APPLICATION NUMBER: 60/039,027
; FILING DATE: February 21, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITH, Guy P.
; REGISTRATION NUMBER: 20,142
; REFERENCE/DOCKET NUMBER: 18950-67 (formerly 18950-62)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 788-5100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: nucleic acid
```

```
;
; STRANDEDNESS: single-stranded DNA
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-861-787A-4

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-861-787A-4 ..
Align seg 1/1 to: US-09-861-787A-4 from: 1 to: 16

1 LysLysAsn 3
|||||
5 AAAAAAAC 13

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-581-970A-69

seq_documentation_block:
; Sequence 69, Application US/09581970A
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY
; APPLICANT: WEBB, DAVID
; APPLICANT: CORYELL, VIRGINIA H.
; APPLICANT: SCHUPP, JAMES M.
; APPLICANT: KEIM, PAUL S.
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS IN SOYBEAN
; FILE REFERENCE: 04-000111US
; CURRENT APPLICATION NUMBER: US/09/581,970A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/US98/26935
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,185
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 798
; SOFTWARE: PatentIn Ver 2.1
; SEQ ID NO: 69
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-581-970A-69

alignment_scores:
  Quality: 16.00 Length: 3
  Ratio: 5.333 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-581-970A-69/rev ..
Align seg 1/1 to reverse of: US-09-581-970A-69 from: 1 to: 16

1 LysLysAsn 3
|||||
9 AAAAAAAC 1

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-941-492-16

seq_documentation_block:
; Sequence 16, Application US/09941492
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd
; APPLICANT: Garcia-Blanco, Mariano M.
; APPLICANT: Puttaraju, Madalah
; APPLICANT: Mansfield, Gary S.
; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN
; SPICEOSOME MEDIATED RNA TRANS-SPICING
```



```
FILE REFERENCE: A31304-BAE (072874.0156)
CURRENT APPLICATION NUMBER: US/09/941,492
CURRENT FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 09/838,858
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 09/756,096
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 09/158,863
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 09/133,717
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 09/087,233
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 08/766,354
PRIOR FILING DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapien
US-09-941-492-16
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-941-492-16/rev ..
```

Align seg 1/1 to reverse of: US-09-941-492-16 from: 1 to: 17

```
1 LysLysAsn 3
|||||
17 AAAAAAAC 6
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-58

```
seq_documentation_block:
; Sequence 58, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID MOI
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING DAT
; TITLE OF INVENTION: THE METHOD
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 17
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-58
```

```
alignment_scores:
  Quality: 16.00      Length: 3
```

```
Ratio: 5.333      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-725-265-58/rev ..

Align seg 1/1 to reverse of: US-09-725-265-58 from: 1 to: 17

```
1 LysLysAsn 3
|||||
17 AAAAAAAC 9
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-59

```
seq_documentation_block:
; Sequence 59, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
; TITLE OF INVENTION: THE METHOD
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 17
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-59
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-528-682-2 x US-09-725-265-59 ..

Align seg 1/1 to: US-09-725-265-59 from: 1 to: 17

```
1 LysLysAsn 3
|||||
17 AAAAAAAC 9
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-60

```
seq_documentation_block:
; Sequence 60, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
```



; TITLE OF INVENTION: THE METHOD  
; FILE REFERENCE: 199953USOXDIV  
; CURRENT APPLICATION NUMBER: US/09/725,265  
; CURRENT FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: US 09/556,127  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: JP 1999-111601  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60  
; LENGTH: 17  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC DNA  
US-09-725-265-60

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-725-265-60 ..  
Align seg 1/1 to: US-09-725-265-60 from: 1 to: 17

1 LysLysAsn 3  
|||||  
1 AAAAAAAC 9

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-63

seq\_documentation\_block:  
; Sequence 63, Application US/09725265  
; GENERAL INFORMATION:  
; APPLICANT: KURANE, RYUICHIRO  
; APPLICANT: KANAGAWA, TAKAHIRO  
; APPLICANT: KANAGAWA, YOICHI  
; APPLICANT: YAMADA, KAZUTAKA  
; APPLICANT: YOKOMAKU, TOYOKAZU  
; APPLICANT: KOYAMA, OSAMU  
; APPLICANT: FURUSHO, KENTA  
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID MOI  
; FILE REFERENCE: 199953USOXDIV  
; CURRENT FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: US/09/725,265  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: JP 1999-111601  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 63  
; LENGTH: 17  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC DNA  
US-09-725-265-63

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-725-265-63/rev ..

Align seg 1/1 to reverse of: US-09-725-265-63 from: 1 to: 17

1 LysLysAsn 3  
|||||  
17 AAAAAAAC 9

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-240

seq\_documentation\_block:  
; Sequence 240, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 240  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-240

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-745-237A-240/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-240 from: 1 to: 17

1 LysLysAsn 3  
|||||  
17 AAGAAGAA 9

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-241

seq\_documentation\_block:  
; Sequence 241, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 241  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-241

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-528-682-2 x US-09-745-237A-241/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-241 from: 1 to: 17



1 LysLysAsn 3  
|||||  
15 AAGAAGAAAT 7

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-242

seq\_documentation\_block:  
; Sequence 242, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 242  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-242

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-745-237A-242/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-242 from: 1 to: 17

1 LysLysAsn 3  
|||||  
14 AAGAAGAAAT 6

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-243

seq\_documentation\_block:  
; Sequence 243, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 243  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-243

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-745-237A-243/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-243 from: 1 to: 17

1 LysLysAsn 3  
|||||  
12 AAGAAGAAAT 4

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-244

seq\_documentation\_block:  
; Sequence 244, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 244  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-244

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-745-237A-244/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-244 from: 1 to: 17

1 LysLysAsn 3  
|||||  
11 AAGAAGAAAT 3

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-245

seq\_documentation\_block:  
; Sequence 245, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease  
; FILE REFERENCE: 400/007 (MBHB00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 245  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-745-237A-245

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-745-237A-245/rev ..

Align seg 1/1 to reverse of: US-09-745-237A-245 from: 1 to: 17

1 LysLysAsn 3  
|||||  
9 AAGAAGAAAT 1

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-745-237A-917



```
seq_documentation_block:
; Sequence 917, Application US/09745237A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: 400/007 (MBHB00-918-A)
; CURRENT APPLICATION NUMBER: US/09745,237A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 4550
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 917
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-745-237A-917
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-745-237A-917/rev ..
Align seg 1/1 to reverse of: US-09-745-237A-917 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
13 AAGAGAAT 5
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-745-237A-918
```

```
seq_documentation_block:
; Sequence 918, Application US/09745237A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: 400/007 (MBHB00-918-A)
; CURRENT APPLICATION NUMBER: US/09745,237A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 4550
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 918
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-745-237A-918
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-745-237A-918/rev ..
Align seg 1/1 to reverse of: US-09-745-237A-918 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
10 AAGAGAAT 2
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-745-237A-1448
```

```
seq_documentation_block:
; Sequence 1448, Application US/09745237A
; GENERAL INFORMATION:
```

```
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: 400/007 (MBHB00-918-A)
; CURRENT APPLICATION NUMBER: US/09745,237A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 4550
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1448
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-745-237A-1448
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-09-745-237A-1448 ..
Align seg 1/1 to: US-09-745-237A-1448 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
2 AAAAAAAC 10
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-043-875-54
```

```
seq_documentation_block:
; Sequence 54, Application US/10043875
; GENERAL INFORMATION:
; APPLICANT: De Smet, Koenraad
; APPLICANT: Stuyver, Lieven
; TITLE OF INVENTION: Method for Detection of Drug-Induced Mutations in the HIV Re
; FILE REFERENCE: 11362-0033-NEPS01 (INNS:033)
; CURRENT APPLICATION NUMBER: US/10/043,875
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/286,102
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: EP 01870085.6
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: EP 01870005.4
; PRIOR FILING DATE: 2001-01-11
; NUMBER OF SEQ ID NOS: 884
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Human Immunodeficiency virus
US-10-043-875-54
```

```
alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-10-043-875-54 ..
Align seg 1/1 to: US-10-043-875-54 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
1 AAGAAAAAT 9
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-1319
```



```
seq_documentation_block:
; Sequence 1319, Application US/10138674
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1319
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-1319
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-10-138-674-1319 ..
```

```
Align seg 1/1 to: US-10-138-674-1319 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
9 AAGAAAGAC 17
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-6287
```

```
seq_documentation_block:
; Sequence 6287, Application US/10138674
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6287
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-6287
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-10-138-674-6287 ..
```

```
Align seg 1/1 to: US-10-138-674-6287 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
9 AAGAAAGAC 17
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-8182
```

```
seq_documentation_block:
; Sequence 8182, Application US/10138674
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8182
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-8182
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-10-138-674-8182 ..
```

```
Align seg 1/1 to: US-10-138-674-8182 from: 1 to: 17
```

```
1 LysLysAsn 3
|||||
2 AAGAAAAAU 10
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-674-8315
```

```
seq_documentation_block:
; Sequence 8315, Application US/10138674
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8315
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-8315
```

```
alignment_scores:
Quality: 16.00 Length: 3
Ratio: 5.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-2 x US-10-138-674-8315 ..
```

```
Align seg 1/1 to: US-10-138-674-8315 from: 1 to: 17
```



```
1 LysLysAsn 3
|||||
2 AAGAAAAAC 10
```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-138-674-8601

```
seq_documentation_block:
; Sequence 8601, Application US/10138674
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138.674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8601
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-8601
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-10-138-674-8601 ..

Align seg 1/1 to: US-10-138-674-8601 from: 1 to: 17

```
1 LysLysAsn 3
|||||
2 AAGAAAAAC 10
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-14

```
seq_documentation_block:
; Sequence 14, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOKAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID MOI
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING DAT
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-14
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-725-265-14 ..

Align seg 1/1 to: US-09-725-265-14 from: 1 to: 18

```
1 LysLysAsn 3
|||||
4 AAAAAAAAT 12
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-15

```
seq_documentation_block:
; Sequence 15, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-15
```

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-725-265-15 ..

Align seg 1/1 to: US-09-725-265-15 from: 1 to: 18

```
1 LysLysAsn 3
|||||
4 AAAAAAAAT 12
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-16

```
seq_documentation_block:
; Sequence 16, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
```



```
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID MOI
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING DAT
; TITLE OF INVENTION: THE METHOD
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-16

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-725-265-16 ..
Align seg 1/1 to: US-09-725-265-16 from: 1 to: 18

1 LysLysAsn 3
|||||
4 AAAAAAAT 12

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-725-265-17

seq_documentation_block:
; Sequence 17, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-17

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-725-265-17 ..
Align seg 1/1 to: US-09-725-265-17 from: 1 to: 18

1 LysLysAsn 3
|||||
4 AAAAAAAT 12

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-725-265-18

seq_documentation_block:
; Sequence 18, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-18

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-725-265-18 ..
Align seg 1/1 to: US-09-725-265-18 from: 1 to: 18

1 LysLysAsn 3
|||||
4 AAAAAAAT 12

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-725-265-19

seq_documentation_block:
; Sequence 19, Application US/09725265
; GENERAL INFORMATION:
; APPLICANT: KURANE, RYUICHIRO
; APPLICANT: KANAGAWA, TAKAHIRO
; APPLICANT: KAMAGATA, YOICHI
; APPLICANT: YAMADA, KAZUTAKA
; APPLICANT: YOKOMAKU, TOYOKAZU
; APPLICANT: KOYAMA, OSAMU
; APPLICANT: FURUSHO, KENTA
; TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING
; FILE REFERENCE: 199953USOXDIV
; CURRENT APPLICATION NUMBER: US/09/725,265
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 09/556,127
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: JP 1999-111601
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 18
; TYPE: DNA
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC DNA
US-09-725-265-19

alignment_scores:
  Quality: 16.00      Length: 3
  Ratio: 5.333       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-2 x US-09-725-265-19 ..
Align seg 1/1 to: US-09-725-265-19 from: 1 to: 18

1 LysLysAsn 3
|||||
4 AAAAAAAT 12
```



FILE REFERENCE: 199953USOXDIV  
CURRENT APPLICATION NUMBER: US/09/725,265  
CURRENT FILING DATE: 2000-11-29  
PRIOR APPLICATION NUMBER: US 09/556,127  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: JP 1999-111601  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 18  
TYPE: DNA  
ORGANISM: ARTIFICIAL SEQUENCE  
FEATURE:  
OTHER INFORMATION: SYNTHETIC DNA  
US-09-725-265-19

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-725-265-19 ..

Align seg 1/1 to: US-09-725-265-19 from: 1 to: 18

1 LysLysasn 3  
|||||||  
4 AAAAAAAT 12

seq\_name: /cgn2\_5/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-725-265-20

seq\_documentation\_block:

Sequence 20, Application US/09725265  
GENERAL INFORMATION:  
APPLICANT: KURANE, RYUICHIRO  
APPLICANT: KANAGAWA, TAKAHIRO  
APPLICANT: KAMAGATA, YOICHI  
APPLICANT: YAMADA, KAZUTAKA  
APPLICANT: YOKOMAKU, TOYOKAZU  
APPLICANT: KOYAMA, OSAMU  
APPLICANT: FURUSHO, KENTA  
TITLE OF INVENTION: METHOD FOR DETERMINING A CONCENTRATION OF TARGET NUCLEIC ACID MO  
TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE METHOD, AND METHOD FOR ANALYZING DAT  
TITLE OF INVENTION: THE METHOD  
FILE REFERENCE: 199953USOXDIV  
CURRENT APPLICATION NUMBER: US/09/725,265  
CURRENT FILING DATE: 2000-11-29  
PRIOR APPLICATION NUMBER: US 09/556,127  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: JP 1999-111601  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20  
LENGTH: 18  
TYPE: DNA  
ORGANISM: ARTIFICIAL SEQUENCE  
FEATURE:  
OTHER INFORMATION: SYNTHETIC DNA  
US-09-725-265-20

alignment\_scores:  
Quality: 16.00 Length: 3  
Ratio: 5.333 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-528-682-2 x US-09-725-265-20 ..

Align seg 1/1 to: US-09-725-265-20 from: 1 to: 18

1 LysLysasn 3  
|||||||  
4 AAAAAAAT 12







OM of: US-09-528-682-3 to: GenEmbl.\* out\_format : pfs

Date: Jun 18, 2002 7:27 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=framet.p2n.model -DEV=xlp  
-Q/cgnt2\_1/USPTO\_spool/US09528682/runat\_18062002\_082443\_7808/app\_query.fasta\_1.689  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LISN=1000 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682.ecgnt1\_8225 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-3

Query length: 107

Database: GenEmbl.\*

Database sequences: 1797656

Database length: 187333701

Search time (sec): 3650.420000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_ba:ECOLITIA	159.00	320.30	1.8e-09	1200	M17894 E.coli heat-labile enterotoxin
gb_hg:AC097953	83.50	117.36	357.17	172040	AC097953 Rattus norvegicus cDNA
gb_vl:AF305738	83.00	165.08	0.7842	697	AF305738 Infectious bursal disease virus
gb_vl:AF305739	83.00	165.08	0.7842	697	AF305739 Infectious bursal disease virus
gb_hg:AC095956	82.50	121.84	200.91	81724	AC095956 Rattus norvegicus cDNA
gb_hg:AC094115	82.50	116.23	412.76	154025	AC094115 Rattus norvegicus cDNA
gb_hg:AC099275	82.50	116.11	419.22	156175	AC099275 Rattus norvegicus cDNA
gb_hg:AC099302	82.50	113.30	601.94	214474	AC099302 Rattus norvegicus cDNA
gb_hg:AC098504	82.50	112.37	677.15	238216	AC098504 Rattus norvegicus cDNA
gb_vl:AF362771	80.50	147.79	7.21	3055	AF362771 Infectious bursal disease virus
gb_hg:AC098193	80.50	115.20	470.84	107518	AC098193 Rattus norvegicus cDNA
gb_hg:AC095291	79.50	111.69	738.95	126055	AC095291 Rattus norvegicus cDNA
gb_hg:AC097044	79.50	105.82	1.6e+03	244530	AC097044 Rattus norvegicus cDNA
gb_ba:AP003587	79.50	102.81	2.3e+03	343550	AP003587 Nostoc sp. PCC 7120
gb_ov:AE008162	79.00	146.57	8.42	2178	AE008162 Xenopus laevis N-CAM
gb_ov:XEUNCAM	79.00	141.83	15.47	3720	M25696 X.laevis neural cell adhesion molecule
gb_hg:AC096608	78.50	111.22	784.09	104705	AC096608 Rattus norvegicus cDNA
gb_hg:AC098098	78.50	106.30	1.5e+03	182686	AC098098 Rattus norvegicus cDNA
gb_hg:AC094978	78.00	108.82	1.1e+03	121958	AC094978 Rattus norvegicus cDNA
gb_ba:AE000673	77.50	127.53	96.90	13097	AE000673 Aquifex aeolicus secA
gb_hg:AC096092	77.50	110.02	914.55	94526	AC096092 Rattus norvegicus cDNA
gb_hg:AC107529	77.50	106.83	1.4e+03	135671	AC107529 Rattus norvegicus cDNA
gb_hg:AC106544	77.50	105.20	1.7e+03	163038	AC106544 Rattus norvegicus cDNA
gb_hg:AC096023	77.50	104.93	1.8e+03	168120	AC096023 Rattus norvegicus cDNA
gb_hg:AC105576	77.50	102.79	2.3e+03	214033	AC105576 Rattus norvegicus cDNA
gb_vl:HVU52889	77.00	146.91	8.07	1303	U52889 Human immunodeficiency virus
gb_vl:HVU52894	77.00	146.91	8.07	1303	U52894 Human immunodeficiency virus
gb_pat:134224	77.00	144.97	10.34	1622	I34224 Sequence 30 from patent
gb_vl:AF443294	77.00	139.37	21.23	3054	AF443294 Infectious bursal disease virus
gb_pat:134221	77.00	138.87	22.62	3230	I34221 Sequence 18 from patent
gb_vl:IBDVPIV	77.00	138.87	22.62	3230	M97346 Infectious bursal disease virus
gb_hg:AC107497	77.00	108.90	1.1e+03	92288	AC107497 Rattus norvegicus cDNA
gb_pi:AF083031	77.00	103.56	2.1e+03	174133	AF083031 Guillardia theta nu
gb_hg:AC087131	77.00	102.61	2.4e+03	193907	AC087131 Mus musculus clone
gb_ba:AP001514	77.00	98.75	3.9e+03	299850	AP001514 Bacillus halodurans
gb_pi:AF132730	76.50	141.14	16.91	2220	AF132730 Homo sapiens unknown
gb_pi:AK027245	76.50	140.39	18.61	2415	AK027245 Homo sapiens cDNA: FLJ
gb_hg:AC094250	76.50	117.29	360.25	32812	AC094250 Rattus norvegicus cDNA
gb_hg:AC095730	76.50	113.12	615.21	52562	AC095730 Rattus norvegicus cDNA
gb_hg:AC095923	76.50	105.30	1.7e+03	127014	AC095923 Rattus norvegicus cDNA

gb_hg:AC106306	76.50	104.23	1.9e+03	143318	AC106306 Rattus norvegicus
gb_hg:AC099071	76.50	99.83	3.4e+03 <td>235713</td> <td>AC099071 Rattus norvegicus</td>	235713	AC099071 Rattus norvegicus
gb_vl:IBDSEGA	76.00	136.98	28.84	3154	D00867 Infectious bursal disease virus
gb_vl:IBDVA	76.00	136.81	29.47	3214	X16107 Chicken infectious bursal disease virus
gb_hg:AC107501	76.00	108.83	1.1e+03 <td>75754</td> <td>AC107501 Rattus norvegicus</td>	75754	AC107501 Rattus norvegicus
gb_hg:AC108886	76.00	102.46	2.4e+03 <td>155423</td> <td>AC108886 Rattus norvegicus</td>	155423	AC108886 Rattus norvegicus
gb_pi:AL136526	76.00	101.17	2.8e+03 <td>179804</td> <td>AL136526 Human DNA sequence</td>	179804	AL136526 Human DNA sequence
gb_hg:AL355680	76.00	100.94	2.9e+03 <td>184654</td> <td>AL355680 Homo sapiens chr</td>	184654	AL355680 Homo sapiens chr
gb_hg:AC094583	76.00	99.96	3.3e+03 <td>206108</td> <td>AC094583 Rattus norvegicus</td>	206108	AC094583 Rattus norvegicus
gb_hg:AC097585	76.00	99.79	3.4e+03 <td>210285</td> <td>AC097585 Rattus norvegicus</td>	210285	AC097585 Rattus norvegicus
gb_sts:CN50615A	75.50	146.12	8.93	997	AL400688 T3 end of clone ASO
gb_in:PRATPI	75.50	129.07	79.50	6838	X65738 P.falciparum gene for
gb_hg:AC095462	75.50	112.62	655.50	43819	AC095462 Rattus norvegicus
gb_hg:AC102241	75.50	108.40	1.1e+03 <td>70550</td> <td>AC102241 Mus musculus clone</td>	70550	AC102241 Mus musculus clone
gb_hg:AC095555	75.50	104.86	1.8e+03 <td>105262</td> <td>AC095555 Rattus norvegicus</td>	105262	AC095555 Rattus norvegicus
gb_hg:AC094595	75.50	102.74	2.3e+03 <td>133693</td> <td>AC094595 Rattus norvegicus</td>	133693	AC094595 Rattus norvegicus
gb_hg:AC097194	75.50	100.33	3.2e+03 <td>175602</td> <td>AC097194 Rattus norvegicus</td>	175602	AC097194 Rattus norvegicus
gb_hg:AC097054	75.50	99.21	3.7e+03 <td>199211</td> <td>AC097054 Rattus norvegicus</td>	199211	AC097054 Rattus norvegicus
gb_vl:AF281227	75.00	148.80	6.33	654	AF281227 Infectious bursal disease virus
gb_vl:IBDVVP2A	75.00	142.44	14.30	1341	X95883 Infectious bursal disease virus
gb_ba:SPU36389	75.00	136.98	28.82	2485	U36389 Synchococcus PCC700
gb_ba:AF305610	75.00	133.34	45.98	3749	AF305610 Borrelia burgdorferi
gb_in:DMTOPII	75.00	127.01	103.57	7664	X61209 D.melanogaster gene
gb_ba:AE006367	75.00	123.11	170.73	11901	AE006367 Lactococcus lactis
gb_hg:AC018322	75.00	106.07	1.5e+03 <td>81532</td> <td>AC018322 Drosophila melan</td>	81532	AC018322 Drosophila melan
gb_in:AC005428	75.00	105.41	1.7e+03 <td>87835</td> <td>AC005428 Drosophila melan</td>	87835	AC005428 Drosophila melan
gb_in:AC099017	75.00	100.22	3.2e+03 <td>157766</td> <td>AC099017 Drosophila melan</td>	157766	AC099017 Drosophila melan
gb_in:AC007082	75.00	99.49	3.5e+03 <td>171375</td> <td>AC007082 Drosophila melan</td>	171375	AC007082 Drosophila melan
gb_hg:AC095105	75.00	98.86	3.8e+03 <td>184001</td> <td>AC095105 Rattus norvegicus</td>	184001	AC095105 Rattus norvegicus
gb_in:AE003663	75.00	94.55	6.7e+03 <td>299449</td> <td>AE003663 Drosophila melan</td>	299449	AE003663 Drosophila melan
gb_sts:CN506KRE	74.50	145.12	10.15	880	AL403248 T7 end of clone AT0
gb_hg:AC094283	74.50	108.40	1.1e+03 <td>55663</td> <td>AC094283 Rattus norvegicus</td>	55663	AC094283 Rattus norvegicus
gb_hg:AC095072	74.50	105.11	1.7e+03 <td>80710</td> <td>AC095072 Rattus norvegicus</td>	80710	AC095072 Rattus norvegicus
gb_hg:AC098996	74.50	103.63	2.1e+03 <td>93531</td> <td>AC098996 Rattus norvegicus</td>	93531	AC098996 Rattus norvegicus
gb_hg:AC098921	74.50	103.54	2.1e+03 <td>96374</td> <td>AC098921 Rattus norvegicus</td>	96374	AC098921 Rattus norvegicus
gb_hg:AC094591	74.50	102.40	2.4e+03 <td>109620</td> <td>AC094591 Rattus norvegicus</td>	109620	AC094591 Rattus norvegicus
gb_hg:AC093970	74.50	99.03	3.7e+03 <td>160371</td> <td>AC093970 Rattus norvegicus</td>	160371	AC093970 Rattus norvegicus
gb_hg:AC098087	74.50	98.89	3.8e+03 <td>162892</td> <td>AC098087 Rattus norvegicus</td>	162892	AC098087 Rattus norvegicus
gb_ba:A7414142	74.50	96.47	5.2e+03 <td>214050</td> <td>A7414142 Versinia pestis</td>	214050	A7414142 Versinia pestis
gb_pat:121030	74.00	132.69	49.96	3180	I21030 Sequence 1 from pate
gb_vl:HVU53046	74.00	98.63	3.9e+03 <td>148892</td> <td>HVU53046 Rattus norvegicus</td>	148892	HVU53046 Rattus norvegicus
gb_vl:HVU52896	73.50	139.56	20.71	1300	U52896 Human immunodeficiency
gb_hg:AC098003	73.50	103.52	2.1e+03 <td>761223</td> <td>AC098003 Rattus norvegicus</td>	761223	AC098003 Rattus norvegicus
gb_hg:AC097409	73.50	100.08	3.3e+03 <td>112233</td> <td>AC097409 Rattus norvegicus</td>	112233	AC097409 Rattus norvegicus
gb_hg:AC094327	73.50	98.88	3.8e+03 <td>124898</td> <td>AC094327 Rattus norvegicus</td>	124898	AC094327 Rattus norvegicus
gb_hg:AC095140	73.50	97.94	4.3e+03 <td>142898</td> <td>AC095140 Rattus norvegicus</td>	142898	AC095140 Rattus norvegicus
gb_hg:AC098497	73.50	96.88	4.9e+03 <td>181042</td> <td>AC098497 Rattus norvegicus</td>	181042	AC098497 Rattus norvegicus
gb_hg:AC103517	73.50	96.66	5.1e+03 <td>165173</td> <td>AC103517 Rattus norvegicus</td>	165173	AC103517 Rattus norvegicus
gb_hg:AC099117	73.50	96.61	5.1e+03 <td>166069</td> <td>AC099117 Rattus norvegicus</td>	166069	AC099117 Rattus norvegicus
gb_hg:AC103174	73.50	96.32	5.3e+03 <td>171655</td> <td>AC103174 Rattus norvegicus</td>	171655	AC103174 Rattus norvegicus
gb_hg:AC095865	73.50	96.13	5.4e+03 <td>175435</td> <td>AC095865 Rattus norvegicus</td>	175435	AC095865 Rattus norvegicus
gb_hg:AC098510	73.50	96.01	5.5e+03 <td>177686</td> <td>AC098510 Rattus norvegicus</td>	177686	AC098510 Rattus norvegicus
gb_hg:AC094413	73.50	95.65	5.8e+03 <td>185150</td> <td>AC094413 Rattus norvegicus</td>	185150	AC094413 Rattus norvegicus
gb_hg:AC097960	73.50	95.17	6.1e+03 <td>195411</td> <td>AC097960 Rattus norvegicus</td>	195411	AC097960 Rattus norvegicus
gb_vl:AF281232	73.00	144.27	11.31	678	AF281232 Infectious bursal disease virus
gb_vl:AF281231	73.00	144.10	11.56	691	AF281231 Infectious bursal disease virus
gb_vl:AF281228	73.00	144.00	11.71	699	AF281228 Infectious bursal disease virus
gb_vl:AF281227	73.00	142.56	14.10	823	AF281227 Infectious bursal disease virus
gb_vl:AF091098	73.00	142.30	14.57	847	AF091098 Infectious bursal disease virus
gb_vl:AF091097	73.00	138.23	24.57	1342	AF363389 Synthetic construct
gb_sy:AF363389	73.00	138.17	24.74	1350	AF464901 Infectious bursal disease virus
gb_vl:AF464901	73.00	138.12	24.74	1350	AF464901 Infectious bursal disease virus
gb_pat:AG6974	73.00	138.09	24.99	1362	AG6974 Sequence 13 from Pat
gb_pat:AT71951	73.00	138.09	24.99	1362	A71951 Sequence 13 from Pat
gb_pat:AR147139	73.00	138.09	24.99	1362	AR147139 Sequence 13 from Pat
gb_pat:BD009827	73.00	138.09	24.99	1362	BD009827 Avian polynucleo
em_pat:EI2060	73.00	138.09	24.99	1362	EI2060 cDNA encoding VP2 pr
gb_vl:AF06701	73.00	137.21	27.88	1350	AF06701 Infectious bursal disease virus
em_pat:EI2069	73.00	137.11	47.34	2391	EI2069 SegA sequence of Inf
gb_vl:IBDSEGA	73.00	131.09	61.35	3004	D00868 Infectious bursal disease virus
gb_vl:AF140705	73.00	131.00	62.10	3036	AF140705 Infectious bursal disease virus
gb_pat:A38328	73.00	130.92	62.72	3063	A38328 Sequence 1 from Pat
gb_pat:A33255	73.00	130.77	63.94	3115	A33255 IBV Edgar strain se
gb_vl:IBDSEGA	73.00	130.56	65.69	3190	D00869 Infectious bursal disease virus







gb_pl:YSCGLN3	71.00	126.83	105.97	3021	M35267 S.cerevisiae nitrogen d	gb_vi:IBDVP523	70.00	124.26	147.32	3183	X92760 Infectious bursal di
gb_vi:AF092171	71.00	126.60	109.17	3101	AF092171 Infectious bursal dis	gb_vi:AF454945	70.00	124.24	147.68	3190	AF454945 Infectious bursal
gb_vi:AF022776	71.00	126.32	113.17	3201	AF362776 Infectious bursal dis	gb_pat:AX074510	70.00	124.05	151.37	3260	AX074510 Sequence 70 from p
gb_pl:SC9379	71.00	103.29	2.2e+03	43100	U18796 Saccharomyces cerevisi	gb_pat:AX138259	70.00	124.05	151.37	3260	AX138259 Sequence 42 from p
gb_pl:AC006587	71.00	97.90	4.3e+03	79262	AC006587 Arabidopsis thaliana	gb_vi:AF240686	70.00	124.05	151.37	3261	AF240686 Infectious bursal
gb_pl:AC0107533	71.00	96.39	5.3e+03	93942	AC107533 Rattus norvegicus cl	gb_vi:IBU0318896	70.00	124.05	151.42	3261	IBU0318896 Infectious bursal
gb_htg:AC105709	71.00	95.83	5.7e+03	100159	AC105709 Rattus norvegicus c	gb_vi:AF092943	70.00	124.03	151.84	3269	AF092943 Infectious bursal
gb_htg:AC096392	71.00	92.27	8.9e+03	149583	AC096392 Rattus norvegicus c	gb_pl:AF092943	70.00	119.58	268.40	5398	M15458 Yeast (S.cerevisiae)
gb_htg:AL3390842	71.00	91.69	9.6e+03	159698	AL3390842 Homo sapiens chromo	gb_pl:SC9379	70.00	119.34	276.89	5548	X03579 Yeast cell division
gb_htg:AC094987	71.00	91.66	9.6e+03	160408	AC094987 Rattus norvegicus c	gb_pl:SC9379	70.00	119.34	276.89	5548	X03579 Yeast cell division
gb_htg:AC094131	71.00	90.40	1.1e+04	184843	AC094131 Rattus norvegicus c	gb_vi:AE002439	70.00	113.53	383.30	10692	L20421 Bacillus circulans b
gb_htg:AL645584	71.00	89.47	1.3e+04	205402	AL645584 Mus musculus chromo	gb_vi:AE002439	70.00	113.53	383.30	10692	L20421 Bacillus circulans b
gb_htg:AF281234	70.50	138.70	23.12	702	AF281234 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_pl:ECU310028	70.50	130.62	65.20	1749	AJ310028 Eriochailus cucullat	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_pl:PN1310053	70.50	130.61	65.20	1750	AJ310053 Paracaula nigricul	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_pat:EO5277	70.50	125.59	124.31	3087	EO5277 cDNA encoding split vi	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF121254	70.50	113.55	581.99	12018	AF121254 Enterococcus faeculi	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097307	70.50	95.07	6.2e+03	96811	AC097307 Rattus norvegicus cl	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095879	70.50	94.57	6.6e+03	102529	AC095879 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC098989	70.50	94.44	6.8e+03	104032	AC098989 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097389	70.50	94.37	6.8e+03	104851	AC097389 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC0973804_2	70.50	93.94	7.2e+03	110000	Continuation (3 of 6) of AC0	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AL356581	70.50	93.48	7.6e+03	115902	AL356581 Homo sapiens chromo	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC094962	70.50	93.47	7.6e+03	115948	AC094962 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095366	70.50	93.40	7.7e+03	116930	AC095366 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095367	70.50	93.37	7.7e+03	117368	AC095367 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095467	70.50	93.34	7.8e+03	117802	AC095467 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095598	70.50	93.21	7.9e+03	119539	AC095598 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095570	70.50	92.04	9.2e+03	136295	AC095570 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC0915618	70.50	91.52	9.8e+03	144605	AC091561 Homo sapiens clone	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AL590643	70.50	91.50	9.8e+03	144865	AL590643 Homo sapiens chromo	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC098608	70.50	91.44	9.9e+03	144964	AC098608 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC094207	70.50	91.43	9.9e+03	146009	AC094207 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_pl:AL137798	70.50	91.43	9.9e+03	146141	AL137798 Human DNA sequence	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095100	70.50	91.38	1.0e+04	146950	AC095100 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095885	70.50	91.32	1.0e+04	147984	AC095885 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097364	70.50	90.90	1.1e+04	155106	AC097364 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095919	70.50	90.63	1.1e+04	159880	AC095919 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AL354666	70.50	90.44	1.1e+04	163290	AL354666 Homo sapiens chromo	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC106662	70.50	90.41	1.1e+04	163974	AC106662 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC103227	70.50	90.41	1.1e+04	164038	AC103227 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC103222	70.50	90.25	1.2e+04	166979	AC103222 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC106247	70.50	90.17	1.2e+04	168429	AC106247 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097820	70.50	90.15	1.2e+04	168882	AC097820 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC106344	70.50	90.04	1.2e+04	170889	AC106344 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC096907	70.50	90.01	1.2e+04	171522	AC096907 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097826	70.50	89.91	1.2e+04	173413	AC097826 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095930	70.50	89.88	1.2e+04	174033	AC095930 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095109	70.50	89.84	1.2e+04	174887	AC095109 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097594	70.50	89.80	1.2e+04	175531	AC097594 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC093961	70.50	89.65	1.2e+04	178655	AC093961 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC096178	70.50	89.49	1.3e+04	181931	AC096178 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC105354	70.50	89.47	1.3e+04	182365	AC105354 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC103013	70.50	89.31	1.3e+04	185713	AC103013 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AL355860	70.50	88.51	1.4e+04	203055	AL355860 Homo sapiens clone	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097948	70.50	88.28	1.5e+04	208587	AC097948 Mus musculus chromo	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC095514	70.50	87.88	1.6e+04	218176	AC095514 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC098938	70.50	87.70	1.6e+04	222687	AC098938 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC097306	70.50	86.88	1.8e+04	244215	AC097306 Mus musculus clone	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC094336	70.50	86.84	1.8e+04	253341	AC094336 Mus musculus clone	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AC103341	70.50	86.63	1.8e+04	251249	AC103341 Rattus norvegicus c	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF281223	70.00	137.88	25.69	684	AF281223 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF281222	70.00	137.84	25.82	687	AF281222 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF281238	70.00	137.74	26.16	695	AF281238 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_pat:AX343661	70.00	133.85	43.07	1078	AX343661 Sequence 1 from Paten	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_vi:AF262030	70.00	131.85	55.66	1351	AF262030 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_vi:AB024076	70.00	131.80	56.04	1359	AB024076 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF281651	70.00	130.47	66.45	1579	AF281651 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF051838	70.00	128.57	84.79	1957	AF051838 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_vi:IBDVKS	70.00	124.67	139.77	3039	I42284 Infectious bursal disea	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF051837	70.00	124.49	143.12	3103	AF051837 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF051837	70.00	124.36	145.48	3148	AF051837 Infectious bursal disea	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi
gb_htg:AF165150	70.00	124.30	146.53	3168	AF165150 Infectious bursal dise	gb_htg:AE007717	70.00	111.05	802.33	14157	AE007717 Clostridium meningi







gb_hhg:AC109525	67.50	96.80	5.0e+03	39025	AC109525	Rattus norvegicus cl	gb_ba:AE005573	67.00	107.40	1.3e+03	10475	AE005573	Escherichia coli
gb_hhg:AC095971	67.50	96.43	5.2e+03	40721	AC095971	Rattus norvegicus cl	gb_ba:AE001513	67.00	105.86	1.6e+03	12456	AE001513	Helicobacter pylori
gb_hhg:AC096279	67.50	94.14	7.0e+03	52725	AC096279	Rattus norvegicus cl	gb_pi:SCX1207K	67.00	101.35	2.8e+03	30473	X75780	S.cerevisiae chromo
gb_hhg:AC097199	67.50	93.39	7.7e+03	57387	AC097199	Rattus norvegicus cl	gb_in:AF024501	67.00	96.35	5.3e+03	38771	AF024501	Caenorhabditis el
gb_hhg:AC107404	67.50	93.37	7.7e+03	57483	AC107404	Rattus norvegicus cl	gb_hhg:AC0112817	67.00	94.40	6.8e+03	43449	AC012817	Drosophila melanog
gb_hhg:AC098192	67.50	93.27	7.8e+03	58187	AC098192	Rattus norvegicus cl	gb_pi:AC011245	67.00	92.47	8.8e+03	57181	AC011245	Homo sapiens BAC
gb_hhg:AC102369	67.50	92.64	8.1e+03	59841	AC102369	Mus musculus clone F	gb_pi:ATW18012	67.00	87.36	1.7e+04	1100678	AL138644	Arabidopsis thal
gb_hhg:AC098446	67.50	92.62	8.5e+03	62843	AC098446	Rattus norvegicus cl	gb_hhg:AC008198	67.00	85.95	2.0e+04	1180678	AC008198	Drosophila melan
gb_hhg:AC099956	67.50	92.55	8.6e+03	63063	AC099956	Mus musculus clone F	gb_hhg:AL645982	67.00	84.88	2.3e+04	133235	AL645982	Mus musculus chr
gb_hhg:AC100935	67.50	92.09	9.1e+03	66445	AC100935	Mus musculus clone F	gb_hhg:AL671961	67.00	83.80	2.6e+04	130393	AL671961	Mus musculus chr
gb_hhg:AC101594	67.50	91.33	1.0e+04	72412	AC101594	Mus musculus clone F	gb_hhg:AC095498	67.00	83.44	2.8e+04	136754	AC095498	Rattus norvegicu
gb_hhg:AC095087	67.50	90.94	1.1e+04	75700	AC095087	Rattus norvegicus cl	gb_hhg:AC095298	67.00	83.40	2.9e+04	157444	AC095298	Rattus norvegicu
gb_pi:AB023033	67.50	90.77	1.1e+04	77129	AB023033	Arabidopsis thaliana	gb_hhg:AC106201	67.00	83.06	2.9e+04	163542	AC106201	Rattus norvegicu
gb_hhg:AC105791	67.50	90.38	1.1e+04	80569	AC105791	Rattus norvegicus cl	gb_hhg:AC106300	67.00	82.97	2.9e+04	165182	AC106300	Rattus norvegicu
gb_hhg:AC096790	67.50	89.44	1.3e+04	89600	AC096790	Rattus norvegicus cl	gb_in:AC017079	67.00	82.90	3.0e+04	166518	AC017079	Drosophila melan
gb_hhg:AC017901	67.50	89.44	1.3e+04	89670	AC017901	Drosophila melanoga	gb_hhg:AC027769	67.00	81.59	3.5e+04	193044	AC027769	Homo sapiens chr
gb_hhg:AC106630	67.50	89.40	1.3e+04	90005	AC106630	Rattus norvegicus cl	gb_hhg:AC080081	67.00	81.22	3.7e+04	201324	AC080081	Homo sapiens chr
gb_hhg:AC103193	67.50	89.36	1.3e+04	90412	AC103193	Rattus norvegicus cl	gb_hhg:AL645807	67.00	81.06	3.7e+04	205077	AL645807	Mus musculus chr
gb_hhg:AC105689	67.50	89.29	1.3e+04	91160	AC105689	Rattus norvegicus cl	gb_ba:AP002086	67.00	80.53	4.0e+04	217594	AP002086	Agrobacterium rh
gb_hhg:AC099348	67.50	88.35	1.5e+04	101368	AC099348	Rattus norvegicus cl	gb_ba:AJ414156	67.00	80.43	4.0e+04	220050	AJ414156	Yersinia pestis
gb_hhg:AC097339	67.50	88.02	1.5e+04	105162	AC097339	Rattus norvegicus cl	gb_pi:AC096776	67.00	80.43	4.0e+04	220119	AC096776	Mus musculus clo
gb_hhg:AC094311	67.50	87.60	1.6e+04	110357	AC094311	Rattus norvegicus cl	gb_hhg:AL611945	67.00	80.07	4.2e+04	239224	AL611945	Mus musculus chr
gb_hhg:AC103260	67.50	87.08	1.7e+04	117062	AC103260	Rattus norvegicus cl	gb_hhg:AC108602	67.00	79.91	4.3e+04	233552	AC108602	Rattus norvegicu
gb_hhg:AC094303	67.50	86.77	1.8e+04	121157	AC094303	Rattus norvegicus cl	gb_hhg:AL645753	67.00	79.68	4.5e+04	239557	AL645753	Mus musculus chr
gb_hhg:AC094733	67.50	85.84	2.0e+04	134533	AC094733	Rattus norvegicus cl	gb_hhg:AL627257	67.00	79.60	4.5e+04	241698	AL627257	Mus musculus chr
gb_hhg:AC094032	67.50	85.60	2.1e+04	138357	AC094032	Rattus norvegicus cl	gb_ba:AP002565	67.00	78.69	5.1e+04	267888	AP002565	Escherichia coli
gb_hhg:AC103488	67.50	85.57	2.1e+04	138784	AC103488	Rattus norvegicus cl	gb_in:AE003554	67.00	78.12	5.4e+04	285752	AE003554	Drosophila melan
gb_hhg:AC094940	67.50	85.39	2.2e+04	141532	AC094940	Rattus norvegicus cl	gb_ba:CU11188X3	67.00	77.19	6.1e+04	337511	CU11188X3	Campylobacter je
gb_hhg:AC091657	67.50	85.32	2.2e+04	142657	AC091657	Rattus norvegicus cl	gb_pi:AY062536	66.50	127.65	95.45	945	AY062536	Arabidopsis thalia
gb_hhg:AC106423	67.50	85.28	2.2e+04	143437	AC106423	Rattus norvegicus cl	gb_pi:AF316564	66.50	120.45	240.09	2129	AF316564	Leptospira interro
gb_hhg:AC095668	67.50	84.79	2.3e+04	151097	AC095668	Rattus norvegicus cl	gb_pi:AF316559	66.50	120.45	240.22	2130	AF316559	Leptospira interro
gb_hhg:AC106459	67.50	84.72	2.3e+04	151548	AC106459	Rattus norvegicus cl	gb_in:AY061160	66.50	120.42	253.72	2235	AY061160	Drosophila melanog
gb_hhg:AC095213	67.50	84.58	2.4e+04	155212	AC095213	Rattus norvegicus cl	gb_pi:ROBKX33	66.50	119.57	268.85	2352	ROBKX33	Bovine rotavirus K33
gb_hhg:AC095565	67.50	84.42	2.4e+04	158078	AC095565	Rattus norvegicus cl	gb_pi:AY062536	66.50	116.50	398.63	3327	AY062536	Arabidopsis thalia
gb_hhg:AC108684	67.50	84.31	2.5e+04	159877	AC108684	Rattus norvegicus cl	gb_pi:AR166375	66.50	116.42	402.58	3356	AR166375	Sequence 4 from pa
gb_hhg:AC095103	67.50	84.08	2.6e+04	162423	AC095103	Rattus norvegicus cl	gb_pi:D86122	66.50	115.83	434.20	3587	D86122	Arabidopsis thaliana
gb_hhg:AC094890	67.50	84.04	2.6e+04	164988	AC094890	Rattus norvegicus cl	em_pi:EL11448	66.50	115.76	438.32	3617	EL11448	cDNA encoding plant
gb_hhg:AC106300	67.50	84.03	2.6e+04	165182	AC106300	Rattus norvegicus cl	gb_pi:AF289053	66.50	115.63	445.76	3671	AF289053	Perina nuda nucleo
gb_hhg:AC092223	67.50	83.69	2.7e+04	171598	AC092223	Drosophila melanoga	gb_in:SHA271051	66.50	115.19	471.64	3858	SHA271051	Schistosoma haemat
gb_hhg:AC106978	67.50	83.66	2.7e+04	172095	AC106978	Rattus norvegicus cl	gb_ba:AB026889	66.50	114.94	486.80	3967	AB026889	Clostridium histol
gb_hhg:AC096437	67.50	83.62	2.7e+04	172437	AC096437	Rattus norvegicus cl	gb_pi:AF394782	66.50	111.97	740.88	5742	AF394782	Homo sapiens rap g
gb_hhg:AC103145	67.50	83.63	2.7e+04	172818	AC103145	Rattus norvegicus cl	gb_ba:AE001283	66.50	111.41	765.14	5914	AE001283	Chlamydia trachom
gb_hhg:AC097612	67.50	83.63	2.7e+04	173469	AC097612	Rattus norvegicus cl	gb_pi:AF117947	66.50	110.70	838.18	6401	AF117947	Homo sapiens PD2 d
gb_hhg:AC098080	67.50	83.59	2.7e+04	173533	AC098080	Rattus norvegicus cl	gb_pi:AF039373	66.50	107.08	1.3e+03	9642	AF039373	Arabidopsis thalia
gb_hhg:AC099604	67.50	83.46	2.8e+04	176163	AC099604	Mus musculus clone	gb_pi:AF039371	66.50	107.08	1.3e+03	9643	AF039371	Arabidopsis thalia
gb_hhg:AC097877	67.50	83.38	2.8e+04	177596	AC097877	Rattus norvegicus cl	gb_pi:AF039376	66.50	106.20	1.5e+03	10640	AF039376	Rattus norvegicu
gb_hhg:AC103049	67.50	83.13	2.9e+04	182795	AC103049	Rattus norvegicus cl	gb_ba:AE007910	66.50	105.85	1.6e+03	11078	AE007910	Agrobacterium tum
gb_pi:AC087599	67.50	83.12	2.9e+04	183046	AC087599	Oryza sativa chromo	gb_ba:AE000963	66.50	104.80	1.8e+03	12471	AE000963	Agrobacterium tum
gb_hhg:AC103275	67.50	83.03	2.9e+04	184922	AC103275	Rattus norvegicus cl	gb_ba:AE001283	66.50	102.40	2.4e+03	16351	AE001283	Chlamydia trachom
gb_hhg:AC084270	67.50	82.92	3.0e+04	187163	AC084270	Rattus norvegicus cl	gb_hhg:AC097864	66.50	99.05	3.7e+03	24651	AC097864	Rattus norvegicu
gb_hhg:AC093967	67.50	82.39	3.2e+04	198694	AC093967	Rattus norvegicus cl	gb_hhg:AC098379	66.50	98.76	3.9e+03	32684	AC098379	Rattus norvegicu
gb_hhg:AC099579	67.50	82.32	3.2e+04	200069	AC099579	Rattus norvegicus cl	gb_hhg:AC094247	66.50	96.27	5.7e+03	32684	AC094247	Rattus norvegicu
gb_hhg:AL626805	67.50	82.32	3.2e+04	200265	AL626805	Mus musculus clone	gb_ba:AF316500	66.50	95.73	5.7e+03	34723	AF316500	Leptospira interr
gb_hhg:AC103338	67.50	82.31	3.2e+04	200159	AC103338	Rattus norvegicus cl	gb_pi:SPAC9	66.50	95.66	5.8e+03	35000	AL121764	S.pombe chromosom
gb_hhg:AC096862	67.50	82.28	3.2e+04	201164	AC096862	Mus musculus chromo	gb_in:CEC05912	66.50	94.04	7.1e+03	42021	Z50872	Caenorhabditis eleg
gb_hhg:AC096862	67.50	82.20	3.2e+04	203032	AC096862	Rattus norvegicus cl	gb_hhg:AC103411	66.50	93.05	8.1e+03	47005	AC103411	Rattus norvegicu
gb_hhg:AC097995	67.50	82.15	3.3e+04	204206	AC097995	Mus musculus clone	gb_hhg:AC109556	66.50	91.13	1.0e+04	58374	AC109556	Mus musculus clon
gb_hhg:AL626806	67.50	81.87	3.4e+04	210827	AL626806	Mus musculus chromo	gb_hhg:AC107477	66.50	90.80	1.1e+04	60383	AC107477	Rattus norvegicu
gb_hhg:AC099295	67.50	81.70	3.5e+04	211712	AC099295	Rattus norvegicus cl	gb_hhg:AC095393	66.50	90.32	1.1e+04	63550	AC095393	Rattus norvegicu
gb_hhg:AC095489	67.50	81.44	3.6e+04	224119	AC095489	Rattus norvegicus cl	gb_hhg:AC105584	66.50	90.28	1.2e+04	64676	AC105584	Rattus norvegicu
gb_hhg:AC073564	67.50	81.19	3.7e+04	227605	AC073564	Mus musculus chromo	gb_hhg:AC096238	66.50	88.62	1.4e+04	77506	AC096238	Rattus norvegicu
gb_hhg:AC096325	67.50	80.97	3.8e+04	233382	AC096325	Rattus norvegicus cl	gb_hhg:AC095945	66.50	88.34	1.5e+04	80008	AC095945	Rattus norvegicu
gb_hhg:AC108602	67.50	80.96	3.8e+04	233552	AC108602	Rattus norvegicus cl	gb_hhg:AC108506	66.50	88.05	1.5e+04	82625	AC108506	Rattus norvegicu
gb_hhg:AC093993	67.50	79.91	4.7e+04	285592	AC093993	Rattus norvegicus cl	gb_hhg:AC102983	66.50	87.74	1.6e+04	85655	AC102983	Rattus norvegicu
gb_hhg:AC091689	67.50	78.18	4.9e+04	294389	AC091689	Rattus norvegicus cl	gb_hhg:AC094686	66.50	87.35	1.7e+04	89515	AC094686	Rattus norvegicu
gb_in:AE003641	67.50	78.15	5.0e+04	299556	AE003641	Drosophila melanoga	gb_hhg:AC095891	66.50	86.86	1.8e+04	94573	AC095891	Rattus norvegicu
gb_in:DROSADH02	67.50	78.15	5.0e+04	320754	AE003408	Drosophila melanoga	gb_hhg:AC096051	66.50	86.75	1.8e+04	95712	AC096051	Rattus norvegicu
gb_ba:NMU81551	67.00	130.40	67.06	780	U81551	Neisseria meningitidis cl	gb_hhg:AC094350	66.50	86.74	1.8e+04	95796	AC094350	Rattus norvegicu
gb_pi:D83985	67.00	118.26	318.01	3071	D83985	Infectious bursal disea	gb_hhg:AC103522	66.50	86.71	1.8e+04	96183	AC103522	Rattus norvegicu
gb_pi:SCNUP100	67.00	114.08	543.50	4923	Z15035	S.cerevisiae gene for	gb_hhg:AC106651	66.50	86.52	1.9e+04	98229	AC106651	Rattus norvegicu
gb_pi:SCVKL068W	67.00	110.99	808.52	6984	Z28068	S.cerevisiae chromosom	gb_hhg:AC098747	66.50	86.51	1.9e+04	98367	AC098747	Rattus norvegicu
gb_ba:LLCPCYRDB	67.00	110.40	871.12	7458	X74207	L.lactis pyrd and pyrF	gb_hhg:AC106556	66.50	86.44	1.9e+04	99184	AC106556	Rattus norvegicu



gb_htg:AC095820	100477	1.9e+04	86.32	66.50	+	AC095820	Rattus norvegicus	gb_htg:AC087875	-	66.50	79.05	4.8e+04	228335	AC087875	Mus musculus
gb_htg:AC094375	105743	2.0e+04	85.87	66.50	-	AC094375	Rattus norvegicus	gb_pr:CNS01DWD	-	66.50	79.04	4.8e+04	228652	AL13128	Human chromosome
gb_htg:AC095679	108813	2.1e+04	85.62	66.50	+	AC095679	Rattus norvegicus	gb_htg:AC099382	+	66.50	78.85	4.9e+04	235258	AC099382	Rattus norvegicus
gb_htg:AC110466	110745	2.1e+04	85.46	66.50	+	AC110466	Rattus norvegicus	gb_ba:AL596170	-	66.50	76.25	6.9e+04	313450	AL596170	Listeria innocua
gb_htg:AC097846	111002	2.1e+04	85.44	66.50	+	AC097846	Rattus norvegicus	gb_vi:AF091099	+	86.00	127.86	92.86	819	AF091099	Infectious bursal d
gb_htg:AC098764	111794	2.2e+04	85.38	66.50	+	AC098764	Rattus norvegicus	gb_vi:AB059304	+	66.00	125.07	132.77	1122	AB059304	Human immunodefici
gb_htg:AC095684	113250	2.2e+04	85.24	66.50	+	AC095684	Rattus norvegicus	gb_vi:AB059306	+	66.00	125.07	132.77	1122	AB059306	Human immunodefici
gb_htg:AC105563	116250	2.3e+04	85.03	66.50	+	AC105563	Rattus norvegicus	gb_pr:HS47014	+	66.00	123.77	156.93	1300	AJ007014	Homo sapiens mrna
gb_htg:OSN00211	116667	2.3e+04	85.00	66.50	+	AL663010	Oryza sativa	gb_pl:SDMATK	+	66.00	122.36	187.99	1524	270186	S. divaricatus chloro
gb_htg:AC093936	120687	2.4e+04	84.70	66.50	+	AC093936	Rattus norvegicus	gb_ro:MMU243485	+	66.00	120.79	229.83	1819	AJ23485	Mus Musculus mrna
gb_htg:AC097887	122056	2.4e+04	84.60	66.50	+	AC097887	Rattus norvegicus	gb_ov:AB008163	+	66.00	119.20	282.00	2178	AB008163	Xenopus laevis N-C
gb_htg:AC106133	129868	2.6e+04	84.05	66.50	+	AC106133	Rattus norvegicus	gb_ov:AB001703	+	66.00	117.25	362.20	2715	AK001703	Homo sapiens cdna
gb_htg:AC093975	130747	2.6e+04	83.99	66.50	+	AC093975	Rattus norvegicus	gb_ov:XEINCA	+	66.00	114.24	532.84	3814	M76710	Xenopus laevis cell
gb_htg:AC099064	134545	2.7e+04	83.74	66.50	+	AC099064	Homo sapiens	gb_pl:AF057708	+	66.00	113.21	608.17	4285	AF057708	Populus balsamifer
gb_htg:AC097075	137140	2.7e+04	83.57	66.50	+	AC097075	Rattus norvegicus	gb_vi:AB006531	+	66.00	106.84	1.4e+03	8797	AB006531	Plautia stali inte
gb_htg:AC103087	137993	2.8e+04	83.40	66.50	+	AC103087	Rattus norvegicus	gb_ba:AE006531	+	66.00	105.36	1.7e+03	10389	AE006541	Streptococcus pyo
gb_htg:AC103009	142434	2.8e+04	83.23	66.50	+	AC103009	Rattus norvegicus	gb_in:AF282241	+	66.00	104.99	1.7e+03	10833	AF282241	Plasmodium falcip
gb_htg:AC095290	151951	3.1e+04	82.66	66.50	+	AC095290	Rattus norvegicus	gb_in:AF282242	+	66.00	104.97	1.7e+03	10863	AF282242	Plasmodium falcip
gb_htg:AC102964	153486	3.1e+04	82.57	66.50	+	AC102964	Rattus norvegicus	gb_in:AF169404	+	66.00	104.92	1.8e+03	10919	AF169404	Plasmodium falcip
gb_htg:AC103123	153625	3.1e+04	82.56	66.50	+	AC103123	Rattus norvegicus	gb_ba:AE000600	+	66.00	103.73	2.0e+03	12489	AE000600	Helicobacter pylor
gb_htg:AC097257	153950	3.1e+04	82.54	66.50	+	AC097257	Rattus norvegicus	gb_ba:AE001778	+	66.00	100.82	3.0e+03	12733	AE001778	Thermotoga mariti
gb_htg:AC094589	154213	3.1e+04	82.53	66.50	+	AC094589	Rattus norvegicus	gb_htg:AC110338	+	66.00	98.32	4.1e+03	23022	AC110338	Rattus norvegicus
gb_htg:AC096059	158224	3.2e+04	82.41	66.50	+	AC096059	Rattus norvegicus	gb_in:CEK07A1	+	66.00	93.63	7.5e+03	39086	Z81097	Caenorhabditis eleg
gb_htg:AC094621	157209	3.2e+04	82.36	66.50	+	AC094621	Rattus norvegicus	gb_htg:AC107108	+	66.00	93.13	8.0e+03	41363	AC107108	Rattus norvegicus
gb_pr:AC074348	157511	3.2e+04	82.34	66.50	+	AC074348	Homo sapiens	gb_pl:SC4357	+	66.00	87.82	1.6e+04	75317	Z38059	S.cerevisiae chromo
gb_htg:AC098991	157734	3.2e+04	82.33	66.50	+	AC098991	Rattus norvegicus	gb_htg:AC096071	+	66.00	87.20	1.7e+04	80780	AC096071	Rattus norvegicus
em_htg:hum:AL353627	158116	3.2e+04	82.31	66.50	+	AL353627	Human DNA	gb_htg:AC098025	+	66.00	86.98	1.8e+04	82865	AC098025	Rattus norvegicus
gb_pr:AC103552	160837	3.3e+04	82.16	66.50	+	AC103552	Homo sapiens	gb_htg:AC091353	+	66.00	86.24	1.9e+04	90047	AC091353	Rattus norvegicus
gb_pr:AP028006	161871	3.3e+04	82.10	66.50	+	AP028006	Homo sapiens	gb_htg:AC095999	+	66.00	86.12	2.0e+04	91295	AC095999	Rattus norvegicus
gb_htg:AC097753	161952	3.3e+04	82.10	66.50	+	AC097753	Rattus norvegicus	gb_htg:AC098899	+	66.00	85.37	2.2e+04	93367	AC098899	Rattus norvegicus
gb_htg:AC105499	162405	3.3e+04	82.07	66.50	+	AC105499	Rattus norvegicus	gb_htg:AC106362	+	66.00	85.23	2.2e+04	100982	AC106362	Rattus norvegicus
gb_htg:OSN00156	162695	3.3e+04	82.05	66.50	+	AL662958	Oryza sativa	gb_htg:AC094955	+	66.00	85.01	2.3e+04	103417	AC094955	Rattus norvegicus
gb_htg:AC103293	163612	3.3e+04	82.00	66.50	+	AC103293	Rattus norvegicus	gb_htg:AL606482	2	66.00	84.47	2.4e+04	110000	Continuation (3 of 5) of	
gb_htg:AC106496	165435	3.4e+04	81.91	66.50	+	AC106496	Rattus norvegicus	gb_htg:AL606482	3	66.00	84.47	2.4e+04	110000	Continuation (4 of 5) of	
gb_htg:AC105834	167109	3.4e+04	81.82	66.50	+	AC105834	Rattus norvegicus	gb_htg:AC094536	+	66.00	83.53	2.7e+04	122225	AC094536	Rattus norvegicus
gb_htg:AC104672	168004	3.4e+04	81.77	66.50	+	AC104672	Rattus norvegicus	gb_pr:PFMA1P1	+	66.00	81.58	3.5e+04	152409	AL03174	Plasmodium falcip
gb_htg:AC094807	168961	3.4e+04	81.72	66.50	+	AC094807	Rattus norvegicus	gb_pr:AC069509	+	66.00	81.56	3.5e+04	152832	AC069509	Homo sapiens
gb_htg:AC094768	168962	3.4e+04	81.69	66.50	+	AC094768	Rattus norvegicus	gb_htg:AP003490	+	66.00	81.34	3.6e+04	156551	AP003490	Oryza sativa chr
AC103002	169494	3.5e+04	81.62	66.50	+	AC103002	Rattus norvegicus	gb_htg:AC079958	+	66.00	80.83	3.8e+04	165865	AC079958	Mus musculus chr
gb_htg:AC098169	170378	3.5e+04	81.65	66.50	+	AC098169	Rattus norvegicus	gb_htg:AC107423	+	66.00	80.73	3.9e+04	167751	AC107423	Homo sapiens chr
gb_htg:AC094714	171046	3.5e+04	81.61	66.50	+	AC094714	Rattus norvegicus	gb_htg:AC099452	+	66.00	80.50	4.0e+04	172199	AC099452	Rattus norvegicus
gb_htg:AC106449	171233	3.5e+04	81.60	66.50	+	AC106449	Rattus norvegicus	gb_htg:AC021180	+	66.00	80.41	4.1e+04	173940	AC021180	Homo sapiens BAC
gb_htg:AC097953	172040	3.5e+04	81.56	66.50	+	AC097953	Rattus norvegicus	gb_htg:AC096593	+	66.00	80.20	4.2e+04	178193	AC096593	Rattus norvegicus
gb_htg:AC094540	173669	3.6e+04	81.48	66.50	+	AC094540	Rattus norvegicus	gb_htg:AC021569	+	66.00	79.46	4.6e+04	193552	AC021569	Homo sapiens chr
AC111720	174393	3.6e+04	81.40	66.50	+	AC111720	Homo sapiens	gb_htg:AC024373	+	66.00	79.27	4.7e+04	197843	AC024373	Homo sapiens chr
gb_htg:AC097821	175218	3.6e+04	81.40	66.50	+	AC097821	Rattus norvegicus	gb_ro:AC083909	+	66.00	78.96	4.9e+04	204857	AC083909	Mus musculus
AC094211	176700	3.6e+04	81.32	66.50	+	AC094211	Rattus norvegicus	gb_htg:AL611937	+	66.00	78.07	5.5e+04	226635	AL611937	Mus musculus chr
AC106087	179303	3.7e+04	81.19	66.50	+	AC106087	Homo sapiens	gb_ro:AL604085	+	66.00	77.02	6.2e+04	254938	AL604085	Mouse DNA sequen
AC103211	179497	3.7e+04	81.18	66.50	+	AC103211	Rattus norvegicus	gb_ba:AP001510	+	66.00	75.52	7.6e+04	302150	AP001510	Bacillus halodur
AC024488	181655	3.7e+04	81.08	66.50	+	AC024488	Homo sapiens	gb_pat:AB5371	-	65.50	130.86	63.17	518	AB5371	Sequence 30 from Pa
AC095154	182794	3.8e+04	81.02	66.50	+	AC095154	Rattus norvegicus	gb_pat:AR154864	-	65.50	130.86	63.17	518	AR154864	Sequence 30 from Pa
AC099182	183271	3.8e+04	81.00	66.50	+	AC099182	Rattus norvegicus	gb_pat:E65389	-	65.50	130.86	63.17	518	E65389	Genome DNA sequen
AC099551	183547	3.8e+04	81.00	66.50	+	AC099551	Homo sapiens	gb_pr:BC007806	-	65.50	129.02	253.88	1763	BC007806	Homo sapiens, clon
AC063929	183624	3.8e+04	80.89	66.50	+	AC063929	Homo sapiens	gb_sy:SNP8GSH	+	65.50	119.40	274.75	1890	D13798	Plasmid pGRS1518-td,
AC096166	186046	3.8e+04	80.87	66.50	+	AC096166	Rattus norvegicus	gb_pr:AK055423	+	65.50	118.45	310.51	2105	AK055423	Homo sapiens cdna
AC099631	186242	3.8e+04	80.86	66.50	+	AC099631	Mus musculus	gb_pat:AX073102	+	65.50	118.09	325.12	2192	AX073102	Sequence 213 from
AC099711	186634	3.8e+04	80.84	66.50	+	AC099711	Mus musculus	gb_pr:AK056296	+	65.50	118.06	326.47	2203	AK056296	Homo sapiens cdna
AC095269	186865	3.8e+04	80.83	66.50	+	AC095269	Rattus norvegicus	gb_in:PFARPHR	+	65.50	117.42	354.08	2260	M86715	P.falciparum 240 kDa
AC094219	188273	3.9e+04	80.76	66.50	+	AC094219	Rattus norvegicus	gb_in:AF303221	+	65.50	117.19	364.99	2427	AF303221	Dictyostelium disc
AC095723	188351	3.9e+04	80.76	66.50	+	AC095723	Rattus norvegicus	gb_in:AF284723	+	65.50	115.94	428.47	2795	AF284763	Plasmodium falcipa
AC095148	190097	3.9e+04	80.68	66.50	+	AC095148	Rattus norvegicus	gb_pl:SCYJL221C	+	65.50	114.27	530.25	3372	249496	S.cerevisiae chromos
AC094541	190132	3.9e+04	80.67	66.50	+	AC094541	Rattus norvegicus	gb_pl:SCYOL157C	+	65.50	112.85	636.63	3961	274899	S.cerevisiae chromos
AC095757	191319	4.0e+04	80.62	66.50	+	AC095757	Rattus norvegicus	gb_pat:A91933	+	65.50	112.00	709.58	4358	AF1933	Sequence 1 from Pate
AC105841	191747	4.0e+04	80.60	66.50	+	AC105841	Rattus norvegicus	gb_pl:AF039372	+	65.50	104.97	1.7e+03	9643	AF039372	Arabidopsis thalia
AC093932	193603	4.0e+04	80.51	66.50	+	AC093932	Rattus norvegicus	gb_htg:AC098338	+	65.50	104.40	1.9e+03	10279	AC098338	Rattus norvegicus
AC098516	193591	4.1e+04	80.42	66.50	+	AC098516	Rattus norvegicus	gb_ba:AE001872	+	65.50	102.91	2.3e+03	12172	AE001872	Beinococcus radio
AC095576	200480	4.2e+04	80.21	66.50	+	AC095576	Rattus norvegicus	gb_ba:AE001091	+	65.50	102.16	2.5e+03	13248	AE001091	Lactococcus ful
AC023173	200970	4.2e+04	80.18	66.50	+	AC023173	Mus musculus	gb_ba:AE006388	+	65.50	101.96	2.6e+03	13549	AE006388	Lactococcus lacti
AC103317	203256	4.2e+04	80.08	66.50	+	AC103317	Rattus norvegicus	gb_htg:AC104177	+	65.50	101.50	2.7e+03	16272	AC104177	Rattus norvegicus
AL671990	205615	4.3e+04	79.98	66.50	+	AL671990	Mus musculus	gb_ba:AF167538	+	65.50	100.33	3.2e+03	16282	AF167538	Trichodesmium sp.
AC094695	208057	4.3e+04	79.88	66.50	+	AC094695	Rattus norvegicus	gb_pl:SC4987	-	65.50	97.94	4.3e+03	21330	Z50178	S.cerevisiae chromo
AC073553															



gb_pl:SC9402	-	65.50	92.60	8.5e+03	38990	I 246921 S.cerevisiae chromosome	gb_htg:AC016649	-	65.50	79.79	4.4e+04	165671	I AC016649 Homo sapiens chr
gb_pl:SCXTEL	-	65.50	92.32	8.9e+03	40257	I 234098 S.cerevisiae chromosome	gb_htg:AC099616	-	65.50	79.75	4.4e+04	166411	I AC099616 Mus musculus clo
gb_in:AF038605	-	65.50	92.37	8.9e+03	40489	I AF038605 Caenorhabditis elegans	gb_htg:AC102398	-	65.50	79.73	4.4e+04	166728	I AC102398 Rattus norvegicus
gb_htg:AC096836	-	65.50	91.30	1.0e+04	45150	I AC096836 Rattus norvegicus cl	gb_htg:AC098945	-	65.50	79.72	4.4e+04	166991	I AC098945 Rattus norvegicus
gb_htg:AC102835	-	65.50	89.68	1.2e+04	54202	I AC102835 Mus musculus clone R	gb_htg:AC096960	-	65.50	79.68	4.5e+04	167698	I AC096960 Rattus norvegicus
gb_htg:AC094578	-	65.50	88.69	1.4e+04	60653	I AC094578 Rattus norvegicus cl	gb_htg:AC095202	-	65.50	79.62	4.5e+04	168827	I AC095202 Rattus norvegicus
gb_htg:AC103637	-	65.50	88.61	1.4e+04	61148	I AC103637 Mus musculus clone R	gb_htg:AC099185	-	65.50	79.58	4.5e+04	169565	I AC099185 Rattus norvegicus
gb_htg:AC103362	-	65.50	88.33	1.5e+04	63136	I AC103362 Mus musculus clone R	gb_htg:AC106664	-	65.50	79.57	4.5e+04	169754	I AC106664 Rattus norvegicus
gb_htg:AC095708	-	65.50	87.32	1.5e+04	63246	I AC095708 Rattus norvegicus cl	gb_htg:AC102986	-	65.50	79.56	4.5e+04	169990	I AC102986 Rattus norvegicus
gb_htg:AC100586	-	65.50	87.50	1.6e+04	66275	I AC100586 Mus musculus clone R	gb_htg:AC095408	-	65.50	79.53	4.5e+04	170138	I AC095408 Rattus norvegicus
gb_pr:AL451009	-	65.50	87.58	1.6e+04	68697	I AL451009 Human DNA sequence	gb_htg:AC096889	-	65.50	79.53	4.5e+04	170638	I AC096889 Rattus norvegicus
gb_htg:AC103489	-	65.50	86.57	1.9e+04	76992	I AC103489 Rattus norvegicus cl	gb_htg:AC106659	-	65.50	79.50	4.6e+04	171155	I AC106659 Rattus norvegicus
gb_htg:AC094328	-	65.50	86.40	1.9e+04	78552	I AC094328 Rattus norvegicus cl	gb_htg:AC094127	-	65.50	79.45	4.6e+04	172027	I AC094127 Rattus norvegicus
gb_htg:AC105512	-	65.50	86.14	2.0e+04	80859	I AC105512 Rattus norvegicus cl	gb_htg:AC105525	-	65.50	79.44	4.6e+04	172391	I AC105525 Rattus norvegicus
gb_htg:AC105859	-	65.50	85.48	2.1e+04	87117	I AC105859 Rattus norvegicus cl	gb_htg:AC106198	-	65.50	79.43	4.6e+04	172601	I AC106198 Rattus norvegicus
gb_in:CEV69H2	-	65.50	85.29	2.2e+04	89004	I 298877 Caenorhabditis elegans	gb_htg:AC106172	-	65.50	79.42	4.6e+04	172687	I AC106172 Rattus norvegicus
gb_htg:AC094113	-	65.50	85.03	2.3e+04	91686	I AC094113 Rattus norvegicus cl	gb_htg:AC095121	-	65.50	79.38	4.6e+04	173584	I AC095121 Rattus norvegicus
gb_htg:AC094864	-	65.50	84.58	2.4e+04	94650	I AC094864 Rattus norvegicus cl	gb_htg:AC096426	-	65.50	79.33	4.7e+04	174434	I AC096426 Rattus norvegicus
gb_htg:AC098628	-	65.50	84.47	2.4e+04	97624	I AC098628 Rattus norvegicus cl	gb_htg:AC080057	-	65.50	79.31	4.7e+04	174879	I AC080057 Homo sapiens chr
gb_htg:AC094625	-	65.50	84.19	2.5e+04	100776	I AC094625 Rattus norvegicus c	gb_htg:AC105356	-	65.50	79.31	4.7e+04	174955	I AC105356 Rattus norvegicus
gb_htg:AC106975	-	65.50	84.12	2.5e+04	101523	I AC106975 Rattus norvegicus c	gb_htg:AC097805	-	65.50	79.21	4.7e+04	176841	I AC097805 Rattus norvegicus
gb_htg:AC106828	-	65.50	83.78	2.6e+04	105588	I AC106828 Rattus norvegicus c	gb_htg:AC005308	-	65.50	79.14	4.8e+04	178273	I AC005308 Plasmodium falci
gb_htg:AC094748	-	65.50	83.77	2.7e+04	105588	I AC094748 Rattus norvegicus c	gb_htg:AC097395	-	65.50	79.07	4.8e+04	178666	I AC097395 Rattus norvegicus
gb_htg:AC105489_2	-	65.50	83.41	2.8e+04	110000	I Continuation (3 of 5) of AC1	gb_htg:AC095623	-	65.50	79.01	4.8e+04	180816	I AC095623 Rattus norvegicus
gb_htg:AC095985	-	65.50	83.09	2.9e+04	114114	I AC095985 Rattus norvegicus c	gb_htg:AC095775	-	65.50	79.00	4.9e+04	181173	I AC095775 Rattus norvegicus
gb_htg:AC095949	-	65.50	83.00	2.9e+04	115214	I AC095949 Rattus norvegicus c	gb_htg:AC055790	-	65.50	78.95	4.9e+04	182087	I AC055790 Homo sapiens chr
gb_htg:AC105724	-	65.50	83.00	2.9e+04	115319	I AC105724 Rattus norvegicus c	gb_htg:AC0955218	-	65.50	78.90	4.9e+04	183219	I AC0955218 Rattus norvegicus
gb_htg:AC094372	-	65.50	82.95	2.9e+04	115923	I AC094372 Rattus norvegicus c	gb_htg:AC094360	-	65.50	78.84	5.0e+04	184432	I AC094360 Rattus norvegicus
gb_htg:AC108090	-	65.50	82.87	3.0e+04	117004	I AC108090 Homo sapiens chromo	gb_htg:AC095459	-	65.50	78.82	5.0e+04	184756	I AC095459 Rattus norvegicus
gb_htg:AC094942	-	65.50	82.74	3.0e+04	118686	I AC094942 Rattus norvegicus c	gb_pr:AC026749	-	65.50	78.82	5.0e+04	184777	I AC026749 Homo sapiens chr
gb_htg:AC106616	-	65.50	82.70	3.1e+04	119308	I AC106616 Rattus norvegicus c	gb_htg:AC105736	-	65.50	78.77	5.0e+04	185839	I AC105736 Rattus norvegicus
gb_htg:AC097437	-	65.50	82.50	3.1e+04	121420	I AC097437 Rattus norvegicus c	gb_htg:AP003714	-	65.50	78.77	5.0e+04	185932	I AP003714 Oryza sativa chr
gb_htg:AC095234	-	65.50	82.26	3.2e+04	132671	I AC095234 Rattus norvegicus c	gb_htg:AC095402	-	65.50	78.63	5.1e+04	18762	I AC095402 Rattus norvegicus
gb_htg:AC026793	-	65.50	81.83	3.4e+04	131613	I AC026793 Homo sapiens chromo	gb_htg:AC095522	-	65.50	78.60	5.1e+04	189362	I AC095522 Rattus norvegicus
gb_htg:AC094293	-	65.50	81.81	3.4e+04	131904	I AC094293 Rattus norvegicus c	gb_htg:AC103441	-	65.50	78.58	5.1e+04	189925	I AC103441 Rattus norvegicus
gb_htg:AC095122	-	65.50	81.56	3.5e+04	135618	I AC095122 Rattus norvegicus c	gb_htg:AL645849	-	65.50	78.02	5.5e+04	203236	I AL645849 Mus musculus chr
gb_htg:AL669829	-	65.50	81.51	3.5e+04	136416	I AL669829 Mus musculus chromo	gb_htg:AC103317	-	65.50	77.98	5.5e+04	203256	I AC103317 Rattus norvegicus
gb_htg:AC094483	-	65.50	81.41	3.6e+04	137944	I AC094483 Rattus norvegicus c	gb_htg:AL669956	-	65.50	77.86	5.6e+04	205879	I AL669956 Mus musculus chr
gb_htg:AC099169	-	65.50	81.40	3.6e+04	138076	I AC099169 Rattus norvegicus c	gb_htg:AC091360	-	65.50	77.67	5.6e+04	210337	I AC091360 Rattus norvegicus
gb_pr:AC010491	-	65.50	81.17	3.7e+04	141706	I AC010491 Homo sapiens chromo	gb_htg:AL663024	-	65.50	77.57	5.8e+04	212915	I AL663024 Mus musculus chr
gb_htg:AC093211	-	65.50	81.16	3.7e+04	141894	I AC093211 Homo sapiens chromo	gb_htg:AL604046	-	65.50	77.36	6.0e+04	217941	I AL604046 Mus musculus chr
gb_htg:AC067790	-	65.50	81.08	3.7e+04	143266	I AC067790 Homo sapiens chromo	gb_htg:AC094290	-	65.50	77.28	6.0e+04	219863	I AC094290 Rattus norvegicus
gb_htg:AP003762	-	65.50	81.02	3.8e+04	144154	I AP003762 Oryza sativa chromo	gb_htg:AC094351	-	65.50	76.89	6.4e+04	229883	I AC094351 Rattus norvegicus
gb_htg:AC099072	-	65.50	81.00	3.8e+04	144420	I AC099072 Rattus norvegicus c	gb_htg:AC096591	-	65.50	76.70	6.5e+04	234893	I AC096591 Mus musculus clo
gb_pr:AP001169	-	65.50	80.86	3.8e+04	146736	I AP001169 Homo sapiens genom	gb_htg:AC096332	-	65.50	75.94	7.2e+04	258816	I AC096332 Rattus norvegicus
gb_htg:AC079759	-	65.50	80.83	3.8e+04	147318	I AC079759 Homo sapiens chromo	gb_htg:AC098521	-	65.50	75.90	7.2e+04	257153	I AC098521 Rattus norvegicus
gb_htg:AC084327	-	65.50	80.78	3.9e+04	148159	I AC084327 Mus musculus chromo	gb_pr:AF239258S3	-	65.50	74.59	8.5e+04	298050	I AF239258 Homo sapiens cal
gb_htg:AC098915	-	65.50	80.77	3.9e+04	148278	I AC098915 Rattus norvegicus c	gb_htg:AC096338	-	65.50	73.93	9.1e+04	321073	I AC096338 Rattus norvegicus
gb_htg:AC096384	-	65.50	80.77	3.9e+04	148314	I AC096384 Rattus norvegicus c	gb_pr:AC096338	-	65.50	73.17	1.0e+05	345999	I AC096338 Staphylococcus a
gb_htg:AC095668	-	65.50	80.60	4.0e+04	151097	I AC095668 Rattus norvegicus c	gb_pl:ACMATK	-	65.00	120.29	245.17	1538	I 270150 A.cathartica chlorop
gb_htg:AC105537	-	65.50	80.58	4.0e+04	151491	I AC105537 Rattus norvegicus c	gb_in:AF072431	-	65.00	118.77	298.08	1803	I AF072431 Dictyostelium disc
gb_htg:AC095392	-	65.50	80.57	4.0e+04	151734	I AC095392 Rattus norvegicus c	gb_pr:AF077019	-	65.00	117.52	338.38	2016	I AF077019 Homo sapiens signa
gb_htg:AC097799	-	65.50	80.53	4.0e+04	152278	I AC097799 Rattus norvegicus c	gb_pr:AF069765	-	65.00	117.52	349.65	2075	I AF069765 Homo sapiens signa
gb_htg:AC096316	-	65.50	80.51	4.0e+04	152718	I AC096316 Rattus norvegicus c	gb_pr:AF069765	-	65.00	116.70	388.56	2277	I AF069765 Homo sapiens signa
gb_htg:AC040919	-	65.50	80.48	4.0e+04	153278	I AC040919 Homo sapiens chromo	gb_om:CFSRP72	-	65.00	113.32	599.15	3334	I 567813 C.familialis SRP72 m
gb_htg:AC105661	-	65.50	80.47	4.0e+04	153476	I AC105661 Rattus norvegicus c	gb_ov:XELPLEX	-	65.00	107.88	1.2e+03	6163	I 381175 Xenopus laevis mRNA
gb_htg:AC106654	-	65.50	80.42	4.1e+04	154210	I AC106654 Rattus norvegicus c	gb_ba:AE006586	-	65.00	102.35	2.4e+03	115707	I AE006586 Streptococcus pyo
gb_htg:AC098261	-	65.50	80.41	4.1e+04	154504	I AC098261 Rattus norvegicus c	gb_ba:BS283337	-	65.00	102.18	2.5e+03	117540	I BS2337 B.subtilis mbl, flh
gb_pr:AL596342	-	65.50	80.38	4.1e+04	154968	I AL596342 Human DNA sequence	gb_ba:AE008596	-	65.00	101.61	2.7e+03	12510	I AE008596 Rickettsia conori
gb_htg:AC105513	-	65.50	80.33	4.1e+04	155907	I AC105513 Rattus norvegicus c	gb_htg:AC109651	-	65.00	100.51	3.1e+03	14172	I AC109651 Rattus norvegicus
gb_htg:AP001845	-	65.50	80.31	4.1e+04	156166	I AP001845 Homo sapiens chromo	gb_ba:STH289861	-	65.00	96.79	5.0e+03	21562	I STH289861 Streptococcus the
gb_htg:AC094654	-	65.50	80.29	4.1e+04	156476	I AC094654 Rattus norvegicus c	gb_htg:AC098323	-	65.00	90.00	1.2e+04	46433	I AC098323 Rattus norvegicus
gb_htg:AC097619	-	65.50	80.25	4.1e+04	157179	I AC097619 Rattus norvegicus c	gb_htg:AC091754	-	65.00	85.69	2.1e+04	75332	I AC091754 Rattus norvegicus
gb_htg:AC098165	-	65.50	80.24	4.1e+04	157388	I AC098165 Rattus norvegicus c	gb_htg:AC095087	-	65.00	85.67	2.1e+04	75700	I AC095087 Rattus norvegicus
gb_htg:AC099218	-	65.50	80.16	4.2e+04	158846	I AC099218 Rattus norvegicus c	gb_htg:AC094515	-	65.00	83.91	2.6e+04	92348	I AC094515 Rattus norvegicus
gb_htg:AC103178	-	65.50	80.16	4.2e+04	158862	I AC103178 Rattus norvegicus c							
gb_htg:AC094205	-	65.50	80.17	4.2e+04	159186	I AC094205 Rattus norvegicus c							
gb_htg:AC106209	-	65.50	80.04	4.2e+04	160479	I AC106209 Rattus norvegicus c							
gb_htg:AC098102	-	65.50	80.03	4.3e+04	161156	I AC098102 Rattus norvegicus c							
gb_pr:AC004774	-	65.50	80.01	4.3e+04	161652	I AC004774 Homo sapiens BAC cl							
gb_htg:AC103237	-	65.50	79.99	4.3e+04	161981	I AC103237 Rattus norvegicus c							
gb_in:AC097292	-	65.50	79.88	4.3e+04	163894	I AC097292 Rattus norvegicus c							
gb_in:PFMAL3P6	-	65.50	79.86	4.4e+04	164399	I 298551 Plasmodium falciparum							



gb\_htg:AC094529 - 65.00 83.62 2.7e+04 95426 ! AC094529 Rattus norvegicus cl  
 gb\_htg:AC098578 - 65.00 82.56 3.1e+04 109575 ! AC098578 Rattus norvegicus c  
 gb\_pr:HSJ556A22 + 65.00 82.43 3.1e+04 109098 ! AL109609 Human DNA sequence  
 gb\_htg:AC097074 + 65.00 82.24 3.2e+04 111556 ! AC097074 Rattus norvegicus c  
 gb\_pr:HSJ599G15 + 65.00 82.15 3.3e+04 112616 ! AL109666 Human DNA sequence  
 gb\_pr:AL138775 - 65.00 81.94 3.4e+04 115304 ! AL138775 Human DNA sequence  
 gb\_htg:AC094609 - 65.00 81.06 3.7e+04 127383 ! AC094609 Rattus norvegicus c  
 gb\_htg:AC080119 - 65.00 79.84 4.4e+04 146375 ! AC080119 Homo sapiens chromo  
 gb\_pr:AL354694 + 65.00 79.42 4.6e+04 153324 ! AL354694 Human DNA sequence  
 gb\_htg:AC016295 + 65.00 79.22 4.7e+04 156906 ! AC016295 Homo sapiens clone  
 gb\_htg:AC098204 + 65.00 79.19 4.7e+04 157340 ! AC098204 Rattus norvegicus c

seq\_name: gb\_ba:ECOLITIIA

seq\_documentation\_block:

LOCUS ECOLITIIA 1200 bp DNA linear BCT 26-APR-1993  
 DEFINITION E.coli heat-labile enterotoxin type IIA (LT-IIa) A and B genes,  
 complete cds.

ACCESSION M17894.1 GI:146671  
 VERSION M17894.1  
 KEYWORDS enterotoxin; heat-labile enterotoxin.  
 SOURCE E.coli (strain SA53) DNA.  
 ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 1200)  
 AUTHORS Pickett,C.L., Weinstein,D.L. and Holmes,R.K.  
 TITLE Genetics of type IIA heat-labile enterotoxin of Escherichia coli:  
 operon fusions, nucleotide sequence, and hybridization studies

JOURNAL J. Bacteriol. 169, 5180-5187 (1987)  
 MEDLINE 88032841  
 FEATURES Location/Qualifiers

1..1200  
 /organism="Escherichia coli"  
 /db\_xref="taxon:562"  
 60..839  
 /note="heat labile enterotoxin type IIA A"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="AA24093.1"  
 /db\_xref="GI:146672"  
 /translation="MKHVLVFFVTSVSANDFRADSRTPDEIRRAGLLPRGQ  
 EAYGPPININLYEHARGVTGNRYNDGVSTVTLRQHLIGQNLGSYNEYIY  
 VVAPFLNIVGLRSPYSPENEFALGGIPLSQIIGWRYVFGAIEGQMQRND  
 YRGLFRLVAPNEDGYOLAGFPNPAWREMPWSTFAPEQCVPNNKFKGVCISA  
 TVNLSKYDLANFKLLKRLALTFNSEDDEFVHGERDEL"  
 829..1200  
 /note="heat labile enterotoxin type IIA B"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="AAA24094.1"  
 /db\_xref="GI:146673"  
 /translation="MSSKKIIGAFVLMTGILSGQVAGVSEHFRNICNTADIVAGV  
 OLKRIADVNTFRGIYVSTGVTGPIGRDYPDNFLSGEIRKTAAMAILSDTKVN  
 LCAKTSPPHIAWELDRS"

BASE COUNT 373 a 175 c 294 g 358 t  
 ORIGIN

alignment\_scores:  
 Quality: 159.00 Length: 220  
 Ratio: 1.807 Gaps: 16  
 Percent Similarity: 40.000 Percent Identity: 31.818

alignment\_block:

US-09-528-682-3 x ECOLITIIA ..

Align seg 1/1 to: ECOLITIIA from: 1 to: 1200

5 ArgAlaLeuGlnGlnAlaTyrGlu.....ProIleGluValAsn... 17

180 AGAGGCGCAGCAGAGCTTATGAGCGCGGAGCCCAATTAACATCAATCT 229

17 ..... 17  
 230 GTATGAGCATGCTCGCGGAACAGTAACGGGAACACACAGATATAATGATG 279  
 18 .....ThrAsnThrValThr..... 22  
 280 GATATGTATCTACAACATGTACAGTTTGAGACAGGTCATTTAATAGGGCAG 329  
 22 ..... 22  
 330 ATATACCTTGGCAGTTAATAATGAATATACATATATATAGTCGCACCAGC 379  
 23 .....GlnIleAsnGlySerAsnGluValProLeuAspGlyA 35  
 380 ACCAAATTTATTGTGATGTGAATGGT.....GTGTTAGGAC 414  
 35 rgTySer.....AsnPheAlaLeuIle..... 42  
 415 GGTATAGTCCATATCCCGAGTGAAATGAATTTGCTGCATTAGGGGAT 464  
 43 .....SerAlaG1 45  
 465 CCCTTATCACAAATTATAGGCTGGTATAGAGTATCTTTTGGTCCGATAGA 514  
 45 uGlyGlyMetGln.....AspGlyAspLeuPheGlyThrV 57  
 515 AGGGGAATGCAGCGAACAACAGGAGATTATCGAGGAGATTATTATTCGAGGT 564  
 57 alaSn.....GlnSer 60  
 565 TAACGGTGCACCTAATGAAGATGGCTATCAACTTGCAGGGTTCCGAGT 614  
 61 AsnPhe.....ProMetSerThrPhe.....GluG1 69  
 615 AACTTCCCGAGCTGGAGAGAATGCCATGGAGTACATTTGCTCCTGAACA 664  
 69 n...ValProAsnAsnLysGluPheLysGly.....ValIleSerA 82  
 665 GTGTGTGCGGAATAATAAAGAAATTTAAAGGAGGGGTGTGCAATTCAGCGA 714  
 82 laAsnVal.....LysTyrAsp...MetAsnPheLysLysLeuLeuArg 95  
 715 CAATGTGCTATCGAAATATGATTTGATGAATTTTAAAAAACCTCTTAAA 764  
 96 .....PheMet...GluAspAspPheIleGlyVa 104  
 765 CCAGCGCTGGCGTTAACGTTTTTCATGACCGAAGATGATTTTATGGTGT 814  
 104 HisGlyGlu 107  
 815 GCATGGAGAA 824

seq\_name: gb\_htg:AC097953

seq\_documentation\_block:

LOCUS AC097953 172040 bp DNA linear HTG 20-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-126M14, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 53 unordered pieces.

ACCESSION AC097953  
 VERSION AC097953.2 GI:17948908  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 172040)

AUTHORS

Muzny,D.N., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,



Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,L., Primus,E., Fu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Somaie,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 172040)  
Worley, K.C.

Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:16327659.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GFRX  
Center clone name: CH230-126M14  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329First call to findPhrapList  
Consensus quality: 148034 bases at least Q40  
Consensus quality: 156693 bases at least Q30  
Consensus quality: 162510 bases at least Q20  
Estimated insert size: 155957; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 2.6x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft.data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft.data.html)).

\* NOTE: This is a 'working draft' sequence. It currently consists of 53 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 8738: contig of 8738 bp in length  
8739 8838: gap of unknown length  
8839 15429: contig of 6591 bp in length

15430  
15530  
24076: contig of 8547 bp in length  
24077  
24176: gap of unknown length  
24177  
28582: contig of 4406 bp in length  
28583  
28682: gap of unknown length  
28683  
34650: contig of 5868 bp in length  
34651  
39278: contig of 4628 bp in length  
39279  
39378: gap of unknown length  
39379  
43355: contig of 3977 bp in length  
43356  
43455: gap of unknown length  
48464: contig of 5009 bp in length  
48465  
48564: gap of unknown length  
48565  
50183: contig of 1619 bp in length  
50184  
50284: gap of unknown length  
50285  
54304: contig of 4021 bp in length  
54305  
54404: gap of unknown length  
59384: contig of 4980 bp in length  
59385  
59484: gap of unknown length  
62418: contig of 2934 bp in length  
62419  
62518: gap of unknown length  
62519  
67293: contig of 4775 bp in length  
67294  
71767: gap of unknown length  
71768  
71867: contig of 4374 bp in length  
76966: contig of 5099 bp in length  
76967  
77066: gap of unknown length  
82885: contig of 5819 bp in length  
82886  
82985: gap of unknown length  
85885: contig of 2900 bp in length  
85886  
85985: gap of unknown length  
88976: contig of 2991 bp in length  
88977  
89076: gap of unknown length  
91424: contig of 2348 bp in length  
91425  
91524: gap of unknown length  
95338: contig of 3914 bp in length  
95339  
95538: gap of unknown length  
97918: contig of 2380 bp in length  
97919  
98018: gap of unknown length  
101817: contig of 3799 bp in length  
101818  
101917: gap of unknown length  
106211: contig of 4294 bp in length  
106311: gap of unknown length  
110097: contig of 3786 bp in length  
110197: gap of unknown length  
110327: contig of 2830 bp in length  
113027: gap of unknown length  
113127: gap of unknown length  
116476: contig of 3349 bp in length  
116576: gap of unknown length  
118469: contig of 1893 bp in length  
118569: gap of unknown length  
121566: contig of 2997 bp in length  
121567  
121666: gap of unknown length  
125119: contig of 3453 bp in length  
125120  
125129: gap of unknown length  
128093: contig of 2874 bp in length  
128094  
128193: gap of unknown length  
128776: contig of 1583 bp in length  
129876: gap of unknown length  
132426: contig of 2550 bp in length  
132526: gap of unknown length  
135110: contig of 2584 bp in length  
135210: gap of unknown length  
137567: contig of 2357 bp in length  
137568  
141734: gap of unknown length  
141735  
141834: gap of unknown length  
143963: contig of 2129 bp in length  
143964  
144063: gap of unknown length  
146871: contig of 2808 bp in length  
146971: gap of unknown length  
148734: contig of 1763 bp in length  
148834: gap of unknown length

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT



```

* 148835 150891: contig of 2057 bp in length
* 150892 150991: gap of unknown length
* 150992 153181: contig of 2190 bp in length
* 153181 153281: gap of unknown length
* 153281 154625: contig of 1344 bp in length
* 154625 154725: gap of unknown length
* 154725 155776: contig of 1051 bp in length
* 155776 155876: gap of unknown length
* 155876 158114: contig of 2338 bp in length
* 158114 158314: gap of unknown length
* 158314 159425: contig of 1111 bp in length
* 159425 159525: gap of unknown length
* 159525 160527: contig of 1002 bp in length
* 160527 160627: gap of unknown length
* 160627 162167: contig of 1540 bp in length
* 162167 162267: gap of unknown length
* 162267 163830: contig of 1563 bp in length
* 163830 163930: gap of unknown length
* 163930 165112: contig of 1182 bp in length
* 165112 165212: gap of unknown length
* 165212 166401: contig of 1189 bp in length
* 166401 166501: gap of unknown length
* 166501 167855: contig of 1154 bp in length
* 167855 167955: gap of unknown length
* 167955 169225: contig of 1470 bp in length
* 169225 169325: gap of unknown length
* 169325 170502: contig of 1177 bp in length
* 170502 170602: gap of unknown length
* 170602 172040: contig of 1438 bp in length.
FEATURES
    source
        1. .172040
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="CH230-126M14"
BASE COUNT  50616 a 31824 c 31989 g 52352 t 5259 others
ORIGIN
alignment_scores
    Quality: 83.50      Length: 138
    Ratio: 1.440       Gaps: 6
    Percent Similarity: 42.029      Percent Identity: 26.087
alignment_block
US-09-528-682-3 x AC097953/rev ..
Align seg 1/1 to reverse of: AC097953 from: 1 to: 172040
21 ValThrGlnIleAsnGlySerAsnGlu.....ValProLeuAs 33
|||||:|||||:|||||:|||||:|||||:|||||:
36521 GTAACCTAACATCAATCTAAATGGAGAGAAACTTGAAGCAATTCACATAA 36472
33 pGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnA 50
|||||:|||||:|||||:|||||:|||||:|||||:
36471 ATCTAGACAGAGCAAT.....CAGACAGCAACGGAGGTCAAAGGG 36431
50 spGlyAspLeuPheGly..... 55
|||:|||||:
36430 ATACGAATTGGAAGGAGGAAATCAAAATATCACTATTTCAGATGATAA 36381
56 .....ThrValAsnG1 59
|||||:
36380 GATAATGCTACTTAAGTCAACCCATAAGTTCCACAGAGAACTACTTAACC 36331
59 nSerAsnPhe..... 62
|||||:
36330 TGATAATTTACGAAAGTGTGGGGGTACAAATTAACATCAACAAATCAG 36281
63 .....PrometSer 65
|||||:
36280 TAGCCTTCCTCTATTCTAAACAGCGCTGAGAAAGAAATTAAGGAATGACA 36231
66 ThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal....IleSe 81
|||:|||||:|||||:|||||:|||||:|||||:

```

```

36230 CCCTTCACAATAGTCCCAATAATATAATAATATCTTGGTGTGCCCTTAAC 36181
81 rAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuLeuArgp 96
: :|||||:|||||:|||||:|||||:|||||:
36180 CAGCAAGTGAAAGATCTGTATGACAAGATTTCAATCTTTGAAGAAG 36131
96 heMetGluAspAsp 100
: :|||||:
36130 AAGTTGAGGAAGAT 36117
seq_name: gb_vi:AF305738
seq_documentation_block:
LOCUS  AF305738                      697 bp  RNA  linear  VRL 25-DEC-2000
DEFINITION  Infectious bursal disease virus isolate 586 VP2 protein gene,
partial cds.
ACCESSION  AF305738
VERSION    AF305738.1  GI:11991622
KEYWORDS   Gumboro virus.
SOURCE     Infectious bursal disease virus
ORGANISM   Viruses: dsRNA viruses; Birnaviridae; Avibirnavirus.
REFERENCE  1 (bases 1 to 697)
AUTHORS    Jackwood,D.J. and Smiley,J.R.
TITLE      Embryo Passed IBDV
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 697)
AUTHORS    Jackwood,D.J. and Smiley,J.R.
TITLE      Direct Submission
JOURNAL    Submitted (13-SEP-2000) Food Animal Health, Ohio State Univ./OARDC,
1680 Madison Ave, Wooster, OH 44691, USA
FEATURES
    source
        1. .697
            /organism="Infectious bursal disease virus"
            /isolate="586"
            /db_xref="taxon:10995"
            /note="chick embryo passage 25"
            <1. .>697
            /note="variable region"
            /codon_start=1
            /product="vp2 protein"
            /protein_id="AAG42307.1"
            /db_xref="GI:11991623"
            /translation="AADDYQFSQYQLGGVTITLFSANIDAITSLSVGGELVFQTSVQ
NIVLGATYILIGFDGTTTTRAVANNGLTAGTNPINFLVFPETNEITQITSIKLE
IVTSKGGGDGDOMSWSAGSLAVTIHGNGYPGALRPVTLVAYERVATGSSVTVAGVS
NFELIPNELAKNLVTEYGRFDPGMNVTKLILSERDLGIKTVTPTREYTDREYEFM
EVADLNSPLKIAGA"
BASE COUNT  193 a 183 c 175 g 146 t
ORIGIN
alignment_scores
    Quality: 83.00      Length: 132
    Ratio: 1.361       Gaps: 7
    Percent Similarity: 46.212      Percent Identity: 23.485
alignment_block
US-09-528-682-3 x AF305738 ..
Align seg 1/1 to: AF305738 from: 1 to: 697
13 ProIleGluValAsn.....ThrAsnThrValThrGlnIleAs 25
|||||:|||||:|||||:|||||:
238 CCTATTCCATTCATCTGTGTTCCCAACCAATCAGATAACCCAG..... 282
25 nGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAlaLeuI 42
|||||:|||||:
283 .....CCAATCATCC.....ATCAAACTGGAGATAG 310
42 leSerAlaGluGlyGlyMetGlnAspGlyAsp..... 52
|||||:|||||:|||||:
311 TAACCTCCAAAAGTGGTGGCCAGATGGGATCAGATGTCATGTCGTGCA 360

```



53 .....LeuPheGlyThrValAsnGlnSerAsnPhePro..... 63  
 ||||| ||||| :|||:|||||  
 361 AGTGGGAGCCTAGCAGTGAACAATCCAGGTGGCAACTATCCAGGGGCCCT 410  
 ||||| :|||:|||||  
 63 ..... 63  
 411 CCGTCAGTCACACTAGTACGCTACGAAAGAGTGGAACACAGGATCTGTGCG 460  
 .....MetSerThrPheGluGlnValProAsnAsnLysGlu 75  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||  
 461 TTACGTCGCTGGGTGAGCAACTCGAGCTGATGCCAAATCTGGAACCTA 510  
 .....MetAsn 89  
 :|||:|||||:|||||:|||||:|||||:|||||  
 511 GCAAAGAACCTGCTTACAGATACGCGCGATTTGACCCAGGAGCTATGAA 560  
 .....MetAsn 89  
 :|||:|||||:|||||:|||||:|||||:|||||  
 89 nPheLysLysLeuLeuArgPheMetGluAspPheIleGlyVal 104  
 :|||:|||||:|||||:|||||:|||||:|||||  
 561 CTACACAAAGTTGATT...CTGAGTGAGAGGACCGCTCTTGGCATC 603

seq\_name: gb\_v1:AF305739

seq\_documentation\_block:  
 LOCUS AF305739 697 bp RNA linear VRL 25-DEC-2000  
 DEFINITION Infectious bursal disease virus isolate 586 VP2 protein gene,  
 partial cds.

ACCESSION AF305739  
 VERSION AF305739.1 GI:11991624

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (15-SEP-2000) Food Animal Health, Ohio State Univ./OARDC,  
 1680 Madison Ave, Wooster, OH 44691, USA

Location/Qualifiers

1. 697

/organism="Infectious bursal disease virus"

/isolate="586"

/db\_xref="taxon:10995"

<1. >697

/note="variable region"

/codon\_start=1

/product="VP2 protein"

/protein\_id="AAG42308.1"

/db\_xref="GI:11991625"

/translation="AADDYQFSQSGVTTTIFSANIDATISLVGGELVFTSVQ  
 NVIGATVILIGFDGTTTTRAVANNGLTAGTDNPIPLVFTPEITOPTITKLE  
 IVTSKSGODGDMWSAGSLAVTIHCGNPGALRPVTLVAYERVATGSVTVVAGVS  
 NFELIPNPELAKNLVTEYCRDPGAMNTKLLSERDLGKTVNPTREYDFREYFM  
 EVADLNPLKKTAGA"

BASE COUNT 193 a 184 c 175 g 145 t

ORIGIN

alignment\_scores:

Quality: 83.00 Length: 132

Ratio: 1.361 Gaps: 7

Percent Similarity: 46.212 Percent Identity: 23.485

alignment\_block:

US-09-528-682-3 x AF305739

Align seq 1/1 to: AF305739 from: 1 to: 697

13 ProileGluValAsn.....ThrAsnThrValThrGlnIleAs 25  
 ||||| ||||| :|||:|||||  
 238 CCTATTCCATCAATCTTGTTCCCAACCAATGAGATACCCAG..... 282  
 25 nGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAlaLeuI 42  
 ||||| :|||:|||||  
 283 .....CCAATCACATCC.....ATCAAACTGGAGATAG 310  
 .....MetSerThrPheGluGlnValProAsnAsnLysGlu 75  
 :|||:|||||:|||||:|||||:|||||:|||||  
 411 CCGTCAGTCACACTAGTACGCTACGAAAGAGTGGAACACAGGATCTGTGCG 460  
 .....MetSerThrPheGluGlnValProAsnAsnLysGlu 75  
 :|||:|||||:|||||:|||||:|||||:|||||  
 461 TTACGTCGCTGGGTGAGCAACTTCGAGCTGATGCCAAATCTGGAACCTA 510  
 .....MetAsn 89  
 :|||:|||||:|||||:|||||:|||||:|||||  
 511 GCAAAGAACCTGCTTACAGATACGCGCGATTTGACCCAGGAGCTATGAA 560  
 .....MetAsn 89  
 :|||:|||||:|||||:|||||:|||||:|||||  
 89 nPheLysLysLeuLeuArgPheMetGluAspPheIleGlyVal 104  
 :|||:|||||:|||||:|||||:|||||:|||||  
 561 CTACACAAAGTTGATT...CTGAGTGAGAGGACCGCTCTTGGCATC 603

seq\_name: gb\_htg:AC095956

seq\_documentation\_block:

LOCUS AC095956

DEFINITION Rattus norvegicus clone CH230-16021, \*\*\* SEQUENCING IN PROGRESS

ACCESSION AC095956

VERSION AC095956.2 GI:17943592

KEYWORDS HTG: HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 81724)

REFERENCE

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbata, J.,

Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J.,

Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,

Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,

Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,

Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,

Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,

Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,

Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,

Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,

Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,

Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,

Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,

Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,

Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,

Hollins, B., Homs, F., Howard, S., Huber, J., Huliy, S., Hume, J.,

Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,

Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,

Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,

Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,

Louise, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,

Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,

Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,

Mei, G., Metzker, M., Miner, Z., Mitchell, T., Mohabbat, K.,

Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,



Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
 Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojokan, I., Rolfe, M.,  
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N.,  
 Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Usmami, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Direct Submission  
 2 (bases 1 to 81724)  
 Worley, K.C.  
 Direct Submission  
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:15627576.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GDZD  
 Center clone name: CH230-16021  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 71574 bases at least Q40  
 Consensus quality: 77370 bases at least Q30  
 Consensus quality: 82981 bases at least Q20  
 Estimated insert size: 62068; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 0.7x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 41 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 5665: contig of 5665 bp in length  
 \* 5666: gap of unknown length  
 \* 5766: contig of 2809 bp in length  
 \* 8575: gap of unknown length  
 \* 8675: gap of unknown length  
 \* 11905: contig of 3231 bp in length  
 \* 11906: gap of unknown length  
 \* 12006: contig of 3543 bp in length  
 \* 15548: gap of unknown length  
 \* 15549: gap of unknown length  
 \* 17549: contig of 2106 bp in length  
 \* 17755: gap of unknown length  
 \* 20363: contig of 2509 bp in length  
 \* 20364: gap of unknown length  
 \* 20464: contig of 2168 bp in length  
 \* 22632: gap of unknown length  
 \* 22732: contig of 1662 bp in length  
 \* 24394: gap of unknown length  
 \* 24494: contig of 2307 bp in length  
 \* 26800: gap of unknown length  
 \* 26801: contig of 3172 bp in length  
 \* 26901: gap of unknown length  
 \* 30073: contig of 2110 bp in length  
 \* 30173: gap of unknown length  
 \* 32282: contig of 2336 bp in length  
 \* 32283: gap of unknown length  
 \* 34718: contig of 2336 bp in length

\* 34719 34818: gap of unknown length  
 \* 34819 37144: contig of 2326 bp in length  
 \* 37145 37244: gap of unknown length  
 \* 37245 38984: contig of 1740 bp in length  
 \* 38985 39084: gap of unknown length  
 \* 39085 40994: contig of 1910 bp in length  
 \* 39085 40994: gap of unknown length  
 \* 40995 43344: contig of 2250 bp in length  
 \* 43345 43444: gap of unknown length  
 \* 43445 44899: contig of 1455 bp in length  
 \* 44900 44999: gap of unknown length  
 \* 45000 46592: contig of 1693 bp in length  
 \* 46593 46792: gap of unknown length  
 \* 46793 48238: contig of 1446 bp in length  
 \* 48239 48338: gap of unknown length  
 \* 48339 50176: contig of 1838 bp in length  
 \* 50177 50276: gap of unknown length  
 \* 50277 51408: contig of 1132 bp in length  
 \* 51409 52793: contig of 1285 bp in length  
 \* 52794 52893: gap of unknown length  
 \* 52894 54196: contig of 1303 bp in length  
 \* 54197 54296: gap of unknown length  
 \* 54297 56017: contig of 1721 bp in length  
 \* 56018 56117: gap of unknown length  
 \* 56118 57818: contig of 1701 bp in length  
 \* 57819 57918: gap of unknown length  
 \* 57919 59916: contig of 1998 bp in length  
 \* 59917 60016: gap of unknown length  
 \* 60017 61273: contig of 1257 bp in length  
 \* 61274 61373: gap of unknown length  
 \* 61374 62810: contig of 1437 bp in length  
 \* 62811 62910: gap of unknown length  
 \* 62911 65433: contig of 2523 bp in length  
 \* 65434 65533: gap of unknown length  
 \* 65534 66773: contig of 1240 bp in length  
 \* 66774 66873: gap of unknown length  
 \* 66874 68013: contig of 1140 bp in length  
 \* 68014 68113: gap of unknown length  
 \* 68114 69332: contig of 1219 bp in length  
 \* 69333 69432: gap of unknown length  
 \* 69433 71004: contig of 1572 bp in length  
 \* 71005 71104: gap of unknown length  
 \* 71105 72526: contig of 1422 bp in length  
 \* 72527 72626: gap of unknown length  
 \* 72627 73972: contig of 1346 bp in length  
 \* 73973 74072: gap of unknown length  
 \* 74073 75110: contig of 1038 bp in length  
 \* 75111 75210: gap of unknown length  
 \* 75211 76565: contig of 1355 bp in length  
 \* 76566 77798: contig of 1133 bp in length  
 \* 77799 77998: gap of unknown length  
 \* 77999 79152: contig of 1254 bp in length  
 \* 79153 79252: gap of unknown length  
 \* 79253 80377: contig of 1125 bp in length  
 \* 80378 80477: gap of unknown length  
 \* 80478 81724: contig of 1247 bp in length.

FEATURES  
 source

1..81724  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-16021"

BASE COUNT 24165 a 15430 c 15170 g 22781 t 4178 others  
 ORIGIN

## alignment\_scores:

Quality: 82.50 Length: 40  
 Ratio: 2.845 Gaps: 2  
 Percent Similarity: 72.500 Percent Identity: 50.000

## alignment\_block:







\* 85053 85152: gap of unknown length  
 \* 85153 contig of 2440 bp in length  
 \* 87593 87592: gap of unknown length  
 \* 87693 91214: contig of 3522 bp in length  
 \* 91215 91314: gap of unknown length  
 \* 91315 93729: contig of 2415 bp in length  
 \* 93730 93829: gap of unknown length  
 \* 93830 96288: contig of 2459 bp in length  
 \* 96289 96388: gap of unknown length  
 \* 96389 98715: contig of 2327 bp in length  
 \* 98716 98815: gap of unknown length  
 \* 98816 101038: contig of 2223 bp in length  
 \* 101039 101138: gap of unknown length  
 \* 101139 103250: contig of 2112 bp in length  
 \* 103251 103250: gap of unknown length  
 \* 103351 106103: contig of 2753 bp in length  
 \* 106104 106203: gap of unknown length  
 \* 106204 107978: contig of 1775 bp in length  
 \* 107979 108078: gap of unknown length  
 \* 108079 109459: contig of 1381 bp in length  
 \* 109460 109559: gap of unknown length  
 \* 109560 111842: contig of 2283 bp in length  
 \* 111843 111942: gap of unknown length  
 \* 111943 114889: gap of unknown length  
 \* 114890 116801: contig of 1812 bp in length  
 \* 116802 116901: gap of unknown length  
 \* 116902 118265: contig of 1364 bp in length  
 \* 118266 118360: gap of unknown length  
 \* 118361 120301: contig of 1935 bp in length  
 \* 120302 120400: gap of unknown length  
 \* 120401 121860: contig of 1460 bp in length  
 \* 121861 121960: gap of unknown length  
 \* 121961 123571: contig of 1611 bp in length  
 \* 123572 123671: gap of unknown length  
 \* 123672 125651: contig of 1980 bp in length  
 \* 125652 125751: gap of unknown length  
 \* 125752 128083: contig of 2332 bp in length  
 \* 128084 128183: gap of unknown length  
 \* 128184 130475: contig of 2292 bp in length  
 \* 130476 130575: gap of unknown length  
 \* 130576 132592: contig of 2017 bp in length  
 \* 132593 132692: gap of unknown length  
 \* 132693 134842: contig of 2150 bp in length  
 \* 134843 134942: gap of unknown length  
 \* 134943 136425: contig of 1483 bp in length  
 \* 136426 136525: gap of unknown length  
 \* 136526 138267: contig of 1742 bp in length  
 \* 138268 138367: gap of unknown length  
 \* 138368 140141: contig of 1774 bp in length  
 \* 140142 140241: gap of unknown length  
 \* 140242 141897: contig of 1656 bp in length  
 \* 141898 141997: gap of unknown length  
 \* 141998 143523: contig of 1526 bp in length  
 \* 143524 143623: gap of unknown length  
 \* 143624 145445: contig of 1822 bp in length  
 \* 145446 145545: gap of unknown length  
 \* 145546 146919: contig of 1374 bp in length  
 \* 146920 147019: gap of unknown length  
 \* 147020 148548: contig of 1529 bp in length  
 \* 148549 148648: gap of unknown length  
 \* 148649 149819: contig of 1171 bp in length  
 \* 149820 149919: gap of unknown length  
 \* 149920 151129: contig of 1210 bp in length  
 \* 151130 151229: gap of unknown length  
 \* 151230 152541: contig of 1312 bp in length  
 \* 152542 152641: gap of unknown length  
 \* 152642 154055: contig of 1414 bp in length.

## FEATURES

Location/Qualifiers  
 1..154055  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-2N6"

## alignment\_scores:

Quality: 82.50 Length: 48  
 Ratio: 2.578 Gaps: 2  
 Percent Similarity: 66.667 Percent Identity: 43.750

## alignment\_block:

US-09-528-682-3 x AC094115 ..

Align seg 1/1 to: AC094115 from: 1 to: 154055

56 ThrValAsnGlnSerAsnPhePromMetSerThrPheGluGlnValProAs 72  
 ||| ::::: ||| |||||  
 72585 ACAAGCCCAAGAAAGAAATAGGAAATGACACCCCTCTCTAATAGTCCCAA 72634

72 nAsnLySgluPheLySglyVal.....IleSerAlaAsnVallys.....T 86

72635 TAAATAAAAAATACCTCGGTGTGACTCTAGCAAGCAAGTGAAGATCTGT 72684

86 yRaspMetAsnPhelysLysLeuLeuArgPheMetGluAspAsp 100

72685 ATGATAATACTTTAAGCCTCTGAAGAAAGAAATTGAAGAAGAC 72728

seq\_name: gb\_htg:AC099275

seq\_documentation\_block:

LOCUS AC099275 156175 bp DNA linear HTG 21-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-32111, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 54 unordered pieces.

## ACCESSION

VERSION AC099275.2 GI:17974747

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 156175)

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,  
 Hollins,B., Homsif,F., Howard,S., Huber,J., Huiy,K.S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Louseghed,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,M., Mawhiney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,M., Nickerson,E., Nwokenwo,S.,  
 Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Fu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,  
 Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,  
 Sisson,I., Sodergren,E., Sonaite,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,



Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wlaczky, R., Woodson, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 156175)  
 Worley, K.C.  
 Direct Submission  
 Submitted (09-NOV-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 21, 2001 this sequence version replaced gi:16874775.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GJOP  
 Center clone name: CH230-32111  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329first call to  
 findPhrapList  
 Consensus quality: 134639 bases at least Q40  
 Consensus quality: 142079 bases at least Q30  
 Consensus quality: 147257 bases at least Q20  
 Estimated insert size: 137844; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 2.3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 54 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 10322: contig of 10322 bp in length  
 \* 10323 10422: gap of unknown length  
 \* 10423 19402: contig of 8980 bp in length  
 \* 19403 19502: gap of unknown length  
 \* 19503 25319: contig of 5817 bp in length  
 \* 25320 25410: gap of unknown length  
 \* 25420 30365: contig of 4946 bp in length  
 \* 30366 30465: gap of unknown length  
 \* 30466 35000: contig of 4535 bp in length  
 \* 35001 35100: gap of unknown length  
 \* 35101 40071: contig of 4971 bp in length  
 \* 40072 40171: gap of unknown length  
 \* 40172 45161: contig of 4990 bp in length  
 \* 45162 45261: gap of unknown length  
 \* 45262 49303: contig of 4042 bp in length  
 \* 49304 49403: gap of unknown length  
 \* 49404 53625: contig of 4222 bp in length  
 \* 53626 53725: gap of unknown length  
 \* 53726 58233: contig of 4508 bp in length  
 \* 58234 58333: gap of unknown length  
 \* 58334 61817: contig of 3484 bp in length  
 \* 61818 61917: gap of unknown length  
 \* 61918 65238: contig of 3321 bp in length  
 \* 65239 65338: gap of unknown length  
 \* 65339 67622: contig of 2284 bp in length  
 \* 67623 67722: gap of unknown length  
 \* 67723 70765: contig of 3043 bp in length  
 \* 70766 70865: gap of unknown length  
 \* 70866 74673: contig of 3808 bp in length  
 \* 74674 74773: gap of unknown length  
 \* 74774 77319: contig of 2546 bp in length  
 \* 77320 77419: gap of unknown length  
 \*  
 \* 77420 80491: contig of 3072 bp in length  
 \* 80492 80591: gap of unknown length  
 \* 80592 83396: contig of 2805 bp in length  
 \* 83397 83496: gap of unknown length  
 \* 83497 85961: contig of 2465 bp in length  
 \* 85962 86061: gap of unknown length  
 \* 86062 87868: contig of 1807 bp in length  
 \* 87869 87968: gap of unknown length  
 \* 87969 91020: contig of 3052 bp in length  
 \* 91021 91120: gap of unknown length  
 \* 91121 93160: contig of 2040 bp in length  
 \* 93161 93260: gap of unknown length  
 \* 93261 94987: contig of 1727 bp in length  
 \* 94988 95087: gap of unknown length  
 \* 95088 97634: contig of 2547 bp in length  
 \* 97635 97734: gap of unknown length  
 \* 97735 99835: contig of 2101 bp in length  
 \* 99836 101751: contig of 1816 bp in length  
 \* 101752 101851: gap of unknown length  
 \* 101852 104752: contig of 2901 bp in length  
 \* 104753 104852: gap of unknown length  
 \* 104853 107784: contig of 2932 bp in length  
 \* 107785 107884: gap of unknown length  
 \* 107885 110476: contig of 2592 bp in length  
 \* 110477 110576: gap of unknown length  
 \* 110577 112349: contig of 1773 bp in length  
 \* 112350 112449: gap of unknown length  
 \* 112450 114160: contig of 1711 bp in length  
 \* 114161 114260: gap of unknown length  
 \* 114261 116230: contig of 1970 bp in length  
 \* 116231 116330: gap of unknown length  
 \* 116331 118217: contig of 1887 bp in length  
 \* 118218 120556: contig of 2239 bp in length  
 \* 120557 120656: gap of unknown length  
 \* 120657 123353: contig of 2697 bp in length  
 \* 123354 123453: gap of unknown length  
 \* 123454 126036: contig of 2583 bp in length  
 \* 126037 126136: gap of unknown length  
 \* 126137 127952: contig of 1816 bp in length  
 \* 127953 128052: gap of unknown length  
 \* 128053 130102: contig of 2050 bp in length  
 \* 130103 130202: gap of unknown length  
 \* 130203 131439: contig of 1237 bp in length  
 \* 131440 131539: gap of unknown length  
 \* 131540 133624: contig of 2085 bp in length  
 \* 133625 133724: gap of unknown length  
 \* 133725 134739: contig of 1015 bp in length  
 \* 134740 134839: gap of unknown length  
 \* 134840 136702: contig of 1863 bp in length  
 \* 136703 136802: gap of unknown length  
 \* 136803 138485: contig of 1883 bp in length  
 \* 138486 138586: gap of unknown length  
 \* 138586 140545: contig of 1960 bp in length  
 \* 140546 140645: gap of unknown length  
 \* 140646 142021: contig of 1376 bp in length  
 \* 142022 142121: gap of unknown length  
 \* 142122 143331: contig of 1210 bp in length  
 \* 143332 143431: gap of unknown length  
 \* 143432 145129: contig of 1698 bp in length  
 \* 145130 145229: gap of unknown length  
 \* 145230 146692: contig of 1463 bp in length  
 \* 146693 146792: gap of unknown length  
 \* 146793 148509: contig of 1717 bp in length  
 \* 148510 148609: gap of unknown length  
 \* 148610 150341: contig of 1732 bp in length  
 \* 150342 150441: gap of unknown length  
 \* 150442 151800: contig of 1359 bp in length  
 \* 151801 151900: gap of unknown length  
 \* 151901 153408: contig of 1508 bp in length  
 \* 153409 153508: gap of unknown length  
 \* 153509 154583: contig of 1075 bp in length

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT







\* 158864 158963: gap of unknown length  
\* 158964 165839: contig of 6876 bp in length  
\* 165840 165939: gap of unknown length  
\* 165940 172366: contig of 6427 bp in length  
\* 172367 172466: gap of unknown length  
\* 172467 177346: contig of 4880 bp in length  
\* 177347 177446: gap of unknown length  
\* 177447 181209: contig of 3763 bp in length  
\* 181210 181309: gap of unknown length  
\* 181310 186947: contig of 5638 bp in length  
\* 186948 187047: gap of unknown length  
\* 187048 193223: contig of 6176 bp in length  
\* 193224 193223: gap of unknown length  
\* 193324 197112: contig of 3789 bp in length  
\* 197113 197212: gap of unknown length  
\* 197213 201092: contig of 3880 bp in length  
\* 201093 201192: gap of unknown length  
\* 201193 204023: contig of 2831 bp in length  
\* 204024 204123: gap of unknown length  
\* 204124 206518: contig of 2395 bp in length  
\* 206519 206618: gap of unknown length  
\* 206619 208071: contig of 1453 bp in length  
\* 208072 208171: gap of unknown length  
\* 208172 210911: contig of 2740 bp in length  
\* 210912 211011: gap of unknown length  
\* 211012 212323: contig of 1312 bp in length  
\* 212324 212423: gap of unknown length  
\* 212424 214474: contig of 2051 bp in length.

#### FEATURES

Location/Qualifiers  
1. .214474  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/chromosome="Mcs1"  
/clone="CH230-132M17"

BASE COUNT 67291 a 42498 c 41787 g 50272 t 2626 others  
ORIGIN

alignment\_scores:  
Quality: 82.50 Length: 42  
Ratio: 2.661 Gaps: 2  
Percent Similarity: 73.810 Percent Identity: 47.619

#### alignment\_block:

US-09-528-682-3 x AC099302 ..

Align seg 1/1 to: AC099302 from: 1 to: 214474

64 MetSerThrpheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79  
||||:||||| ||||| ||||| ||||| ||||| |||||  
91377 ATGACAACTTCATACAGTCCCAATAATATAATACCTTGGTGAC 91426

80 .IleSerAlaAsnValLys.....TyrAspMetAsnPhelLysLeuL 94  
||||:||||| ||||| ||||| ||||| ||||| |||||  
91427 TTTAAACGAAGCAAGTGAAGATCTGATGATGAAGAACTTCAAGTCTGA 91476

94 euArgPheMetGluAspPheile 102

91477 AGAAAGAAATGGAAGAGATCTTATA 91502

seq\_name: gb\_htg:AC098504

seq\_documentation\_block:

LOCUS AC098504 238216 bp DNA linear HTG 20-DEC-2001  
DEFINITION Rattus norvegicus chromosome Mcs1 clone CH230-40F5, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 71 unordered pieces.

ACCESSION AC098504

VERSION AC098504.2 GI:17973142

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

#### Rattus.

##### REFERENCE AUTHORS

1 (bases 1 to 238216)  
Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen C.,  
Alsbrooks S.L., Amaratunge H.C., Are J.R., Banks T., Barbaria J.,  
Benton J., Bivaga K., Blankenburg K., Bonnin D., Bouck J.,  
Bowie S., Brieva M., Brown M., Bryant N.P., Buhay C.,  
Burch P., Burkett C., Burrell K.L., Byrd N.C., Carron T.F.,  
Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,  
Chen Z., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C.,  
Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C.,  
Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O.,  
Denn A.L., Ding Y., Dinh H.H., Douthwaite K.J., Draper H.,  
Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,  
Elhaj C., Escotto P., Falls T., Ferraguto D., Flagg N., Ford J.,  
Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,  
Garza N., Gill R., Gorrell J.H., Guevara W., Gunaratne P., Hale S.,  
Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hawes A.,  
Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C.,  
Hollins B., Homs J.F., Howard S., Huber J., Hulyk S., Hume J.,  
Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S.,  
Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J.,  
Kovar C., Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C.,  
Lewis L., Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W.,  
Louiseged H., Lozano R.J., Lu X., Lucier A., Lucier R., Luna R.,  
Ma J., Maheshwari M., Mapa P., Martin R., Martindale A.,  
Martinez E., Massey E., Mawhney E., McLeod M.P., Meador M.,  
Mei G., Metzker M., Miner G., Miner Z., Mitchell T., Mohabbat K.,  
Morgan M., Morris S., Moser M., Neal D., Newton J., Newton N.,  
Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokenkwo S.,  
Ogulu M., Okwuonu G., Oragunye N., Oviedo R., Pace A., Payton B.,  
Peirys J., Perez L., Peters L., Pickens R., Primus E., Pu L.L.,  
Quiles M., Ren Y., Rives M., Rojas A., Rojibokan I., Rolfe M.,  
Ruiz S., Savery G., Scherer S., Scott G., Shen G., Shostari N.,  
Sisson I., Sodergren E., Sonaike T., Sparks A., Stanley H.,  
Stone H., Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K.,  
Tang H., Tansey J., Taylor C., Taylor T., Telford B., Thomas N.,  
Thomas S., Usmani K., Vasquez L., Vera V., Villalon D., Vinson R.,  
Wall R., Wang S., Ward-Moore S., Warren R., Washington C.,  
Watlington S., Williams G., Williamson A., Wlezyk R., Wooden S.,  
Worley K., Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,  
Weinstock G. and Gibbs R.

##### TITLE JOURNAL

Unpublished  
2 (bases 1 to 238216)

##### REFERENCE AUTHORS

Worley K.C.

##### TITLE JOURNAL

Submitted (24-OCT-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:16356709.

##### COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GHFH  
Center clone name: CH230-40F5  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329First call to  
findPhrapList  
Consensus quality: 206528 bases at least Q40  
Consensus quality: 217991 bases at least Q30  
Consensus quality: 227959 bases at least Q20  
Estimated insert size: 210373; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 2.3x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 71 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is



\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

1 13234: contig of 13234 bp in length  
\* 13235 13334: gap of unknown length  
\* 13335 23485: contig of 10151 bp in length  
\* 23486 23585: gap of unknown length  
\* 23586 33096: contig of 9511 bp in length  
\* 33097 33196: gap of unknown length  
\* 33197 40268: contig of 7072 bp in length  
\* 40269 40368: gap of unknown length  
\* 40369 48964: contig of 8596 bp in length  
\* 48965 49064: gap of unknown length  
\* 49065 54808: contig of 5744 bp in length  
\* 54809 54908: gap of unknown length  
\* 54909 65294: contig of 10386 bp in length  
\* 65295 65394: gap of unknown length  
\* 65395 72516: contig of 7122 bp in length  
\* 72517 72616: gap of unknown length  
\* 72617 79479: contig of 6863 bp in length  
\* 79480 79579: gap of unknown length  
\* 79580 86800: contig of 7221 bp in length  
\* 86801 86900: gap of unknown length  
\* 86901 93042: contig of 6142 bp in length  
\* 93043 93142: gap of unknown length  
\* 93143 100155: contig of 7013 bp in length  
\* 100156 100255: gap of unknown length  
\* 100256 104941: contig of 4686 bp in length  
\* 104942 105041: gap of unknown length  
\* 105042 107963: contig of 2922 bp in length  
\* 107964 108063: gap of unknown length  
\* 108064 113675: contig of 5612 bp in length  
\* 113676 113775: gap of unknown length  
\* 113776 117901: contig of 4126 bp in length  
\* 117902 118001: gap of unknown length  
\* 118002 121900: contig of 3899 bp in length  
\* 121901 122000: gap of unknown length  
\* 122001 126077: contig of 4077 bp in length  
\* 126078 126177: gap of unknown length  
\* 126178 130699: contig of 4522 bp in length  
\* 130700 130799: gap of unknown length  
\* 130800 133531: contig of 2732 bp in length  
\* 133532 133631: gap of unknown length  
\* 133632 137491: contig of 3860 bp in length  
\* 137492 137591: gap of unknown length  
\* 137592 139973: contig of 2382 bp in length  
\* 139974 140073: gap of unknown length  
\* 140074 143545: contig of 3472 bp in length  
\* 143546 143645: gap of unknown length  
\* 143646 147633: contig of 3988 bp in length  
\* 147634 147733: gap of unknown length  
\* 147734 151739: contig of 4006 bp in length  
\* 151740 151839: gap of unknown length  
\* 151840 154469: contig of 2630 bp in length  
\* 154470 154569: gap of unknown length  
\* 154570 158637: contig of 4068 bp in length  
\* 158638 158737: gap of unknown length  
\* 158738 162043: contig of 3306 bp in length  
\* 162044 162143: gap of unknown length  
\* 162144 164784: contig of 2641 bp in length  
\* 164785 164884: gap of unknown length  
\* 164885 166992: contig of 2108 bp in length  
\* 166993 167092: gap of unknown length  
\* 167093 168942: contig of 1850 bp in length  
\* 168943 169042: gap of unknown length  
\* 169043 170705: contig of 1663 bp in length  
\* 170706 170805: gap of unknown length  
\* 170806 172524: contig of 1719 bp in length  
\* 172525 172624: gap of unknown length  
\* 172625 175894: contig of 3270 bp in length  
\* 175895 175994: gap of unknown length

\* 175995 177334: contig of 1340 bp in length  
\* 177335 177434: gap of unknown length  
\* 177435 179248: contig of 1814 bp in length  
\* 179249 179348: gap of unknown length  
\* 179349 181293: contig of 1945 bp in length  
\* 181294 181393: gap of unknown length  
\* 181394 183249: contig of 1856 bp in length  
\* 183250 183349: gap of unknown length  
\* 183350 185855: contig of 2506 bp in length  
\* 185856 185955: gap of unknown length  
\* 185956 187709: contig of 1754 bp in length  
\* 187710 187809: gap of unknown length  
\* 187810 190193: contig of 2384 bp in length  
\* 190194 190293: gap of unknown length  
\* 190294 191889: contig of 1696 bp in length  
\* 191890 192089: gap of unknown length  
\* 192090 193918: contig of 1829 bp in length  
\* 193919 194018: gap of unknown length  
\* 194019 195420: contig of 1402 bp in length  
\* 195421 195520: gap of unknown length  
\* 195521 197552: contig of 2032 bp in length  
\* 197553 197652: gap of unknown length  
\* 197653 199345: contig of 1693 bp in length  
\* 199346 199445: gap of unknown length  
\* 199446 201141: contig of 1696 bp in length  
\* 201142 201241: gap of unknown length  
\* 201242 202754: contig of 1513 bp in length  
\* 202755 202854: gap of unknown length  
\* 202855 204421: contig of 1567 bp in length  
\* 204422 204521: gap of unknown length  
\* 204522 206255: contig of 1734 bp in length  
\* 206256 206355: gap of unknown length  
\* 206356 208503: contig of 2148 bp in length  
\* 208504 208603: gap of unknown length  
\* 208604 209943: contig of 1340 bp in length  
\* 209944 210043: gap of unknown length  
\* 210044 211437: contig of 1394 bp in length  
\* 211438 211537: gap of unknown length  
\* 211538 213188: contig of 1651 bp in length  
\* 213189 213288: gap of unknown length  
\* 213289 215298: contig of 2010 bp in length  
\* 215299 215398: gap of unknown length  
\* 215399 217318: contig of 1920 bp in length  
\* 217319 217418: gap of unknown length

alignment\_scores:           Quality: 82.50           Length: 42  
                              Ratio: 2.661           Gaps: 2  
Percent Similarity: 73.810   Percent Identity: 47.619

alignment\_block:  
US-09-528-682-3 x AC098504/rev ..

Align seg 1/1 to reverse of: AC098504 from: 1 to: 238216

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal... 79  
|||||  
202040 ATGCAACCTTCATACAGTCCCAATAATAATAAAATACCTTGGTGTGAC 201991  
|||||

80 .IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94  
|||||  
201990 TTTAACCAAGCAAGTGAAGATCTGTATGATAAGAACTTCAAGTCTCTGA 201941  
|||||

94 euArgPheMetGluAspPhePhe 102  
|||

201940 AGAAGAAATTCAGAGAGATCTTATA 201915  
|||

seq\_name: gb\_vi:AF362771

seq\_documentation\_block:  
LOCUS AF362771  
DEFINITION Infectious bursal disease virus strain Cu-1 M segment A structural  
polyprotein gene, partial cds.

3055 bp   RNA   linear   VRL 02-JUL-2001



```

57 ValAsnGlnSerAsnPhePro..... 63
   ::::||| ::::|||::|||
1086 ATCAATAGTGTGGCAACTATCCAGGGGCCCTCGTCGCCGTACGCTAGTGGC 1135
      .....
64 ..... MetSert 66
      ::::|||
1136 CTACGAAGAAGTGGCAACAGGATCCGCTTACGGTCGCTGGGGTGAGCA 1185
      .....
66 hrPheGluInValProAsnAsnLysGluPheLysGlyValIleSerAla 82
   ::::||| ::::||| ::::|||::|||::|||
1186 ACTTCGAGCTGATCCCAAATCCTGAACTAGCAAAGAACCCTGGTTACAGAA 1235
      .....
83 AsnValLysTyrrAsp.....MetAsnPheLysLysLeuLeuargPh 96
   ::::||| ::::|||::|||::|||::|||::|||::|||::|||::|||
1236 TACGGCCGATTGACCCAGGAGCATTAACATAACAAAATTGATA...CT 1282
      .....
96 eMetGluAspAspPheIleGlyVal 104
   : ||| ||| ::::|||::|||::|||::|||
1283 GAGTGAGAGGGACCGCTCTTGGCATC 1307

seq_name: gb_htr:AC098193

seq_documentation_block:
LOCUS AC098193
DEFINITION Rattus norvegicus clone CH230-56k11, *** SEQUENCING:
***, 63 unordered pieces.
ACCESSION AC098193
VERSION AC098193.4 GI:17973703
KEYWORDS HTG; HTGS, PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Rattus.
REFERENCE 1 (bases 1 to 107518)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R.,

```

Alsbrooks, S. L., Amaraturunge, H. C., Are, J. R., Banks, T., Barbarella, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Bunay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Deng, A. L., Ding, Y., Dinsh, H. H., Douthwaite, K. J., Draper, H., Dunn-Rocha, S., Durbin, K. J., Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Joliviet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovac, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewtas, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisgeed, H., Lozardo, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Mosser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonuo, G., Orangunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savory, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villallon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyka, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.



```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
-----
Wainstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 107518)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064560.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRHW
Center clone name: CH230-56K11
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 100954 bases at least Q40
Consensus quality: 112766 bases at least Q30
Consensus quality: 120930 bases at least Q20
Estimated insert size: 60056; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 0.7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 6151: contig of 6150 bp in length
* 6250: gap of unknown length
* 9651: contig of 3414 bp in length
* 9664: gap of unknown length
* 9764: gap of unknown length
* 9765: contig of 4369 bp in length
* 14133: contig of 2509 bp in length
* 14233: gap of unknown length
* 14234: contig of 16742 bp in length
* 16743: gap of unknown length
* 16744: contig of 1192 bp in length
* 16843: contig of 1192 bp in length
* 18035: gap of unknown length
* 18134: contig of 1371 bp in length
* 18135: gap of unknown length
* 19505: gap of unknown length
* 19506: contig of 3377 bp in length
* 22983: gap of unknown length
* 22983: contig of 1704 bp in length
* 23083: gap of unknown length
* 24787: contig of 1704 bp in length
* 24787: gap of unknown length
* 24887: contig of 2203 bp in length
* 27090: gap of unknown length
* 27189: contig of 2498 bp in length
* 29687: gap of unknown length
* 29688: contig of 1825 bp in length
* 31612: gap of unknown length
* 31613: contig of 1618 bp in length
* 31713: contig of 1618 bp in length
* 33330: gap of unknown length
* 33430: contig of 1167 bp in length
* 33431: gap of unknown length
* 34597: gap of unknown length
* 34598: contig of 1958 bp in length
* 34698: gap of unknown length
* 36555: gap of unknown length
* 36556: contig of 1346 bp in length
* 36756: gap of unknown length
* 38101: contig of 1840 bp in length
* 38201: gap of unknown length
* 40041: contig of 1840 bp in length
* 40042: gap of unknown length
* 40141: contig of 1817 bp in length
* 41958: gap of unknown length
* 41959: contig of 1111 bp in length
* 42059: gap of unknown length
*
* 43170: gap of unknown length
* 43270: contig of 1538 bp in length
* 44807: gap of unknown length
* 44808: contig of 2391 bp in length
* 47298: gap of unknown length
* 47398: contig of 1996 bp in length
* 49394: gap of unknown length
* 49494: contig of 1564 bp in length
* 51058: gap of unknown length
* 51059: contig of 1012 bp in length
* 52170: gap of unknown length
* 52171: contig of 1398 bp in length
* 53668: gap of unknown length
* 53769: contig of 1205 bp in length
* 54974: gap of unknown length
* 55073: contig of 1316 bp in length
* 56389: gap of unknown length
* 56489: contig of 1649 bp in length
* 58139: gap of unknown length
* 58238: contig of 1131 bp in length
* 59369: gap of unknown length
* 59469: contig of 1557 bp in length
* 61026: gap of unknown length
* 61226: contig of 1012 bp in length
* 62138: gap of unknown length
* 62238: contig of 1279 bp in length
* 63517: gap of unknown length
* 64662: contig of 1045 bp in length
* 64762: gap of unknown length
* 66002: contig of 1240 bp in length
* 66102: gap of unknown length
* 67114: contig of 1012 bp in length
* 67214: gap of unknown length
* 68299: contig of 1715 bp in length
* 69029: gap of unknown length
* 70225: contig of 1196 bp in length
* 70325: gap of unknown length
* 71636: contig of 1311 bp in length
* 71736: gap of unknown length
* 72824: contig of 1088 bp in length
* 72924: gap of unknown length
* 74107: contig of 1183 bp in length
* 74207: gap of unknown length
* 75740: contig of 1533 bp in length
* 75840: gap of unknown length
* 77352: contig of 1512 bp in length
* 77452: gap of unknown length
* 79422: contig of 1970 bp in length
* 79522: gap of unknown length
* 80997: contig of 1475 bp in length
* 81097: gap of unknown length
* 82320: contig of 1223 bp in length
* 82420: gap of unknown length
* 83782: contig of 1362 bp in length
* 83882: gap of unknown length
* 84910: contig of 1028 bp in length
* 85010: gap of unknown length
* 85011: contig of 1561 bp in length
* 86571: gap of unknown length
* 86572: contig of 1061 bp in length
* 87732: gap of unknown length
* 87832: contig of 1368 bp in length
* 89200: gap of unknown length
* 89300: gap of unknown length
* 90489: contig of 1189 bp in length
* 90589: gap of unknown length
* 91849: contig of 1260 bp in length
* 91949: gap of unknown length
* 93403: contig of 1454 bp in length
* 93503: gap of unknown length
* 94523: contig of 1020 bp in length
* 94623: gap of unknown length
* 95755: contig of 1132 bp in length
* 95855: gap of unknown length

```







\* 72635 74993: contig of 2359 bp in length  
 \* 74994 75093: gap of unknown length  
 \* 75094 77443: contig of 2350 bp in length  
 \* 77444 77543: gap of unknown length  
 \* 77544 79010: contig of 1467 bp in length  
 \* 79011 79110: gap of unknown length  
 \* 81179 81679: contig of 2569 bp in length  
 \* 81680 81779: gap of unknown length  
 \* 81780 84606: contig of 2827 bp in length  
 \* 84607 84706: gap of unknown length  
 \* 84707 87065: contig of 2359 bp in length  
 \* 87066 87165: gap of unknown length  
 \* 87166 89890: contig of 2725 bp in length  
 \* 89891 89990: gap of unknown length  
 \* 89991 92950: contig of 2960 bp in length  
 \* 92951 93050: gap of unknown length  
 \* 93051 95184: contig of 2134 bp in length  
 \* 95185 95284: gap of unknown length  
 \* 95285 96730: contig of 1446 bp in length  
 \* 96731 96830: gap of unknown length  
 \* 96831 98927: contig of 2097 bp in length  
 \* 98928 101350: contig of 2323 bp in length  
 \* 101351 101450: gap of unknown length  
 \* 101451 103552: contig of 2102 bp in length  
 \* 103553 103653: gap of unknown length  
 \* 103653 105349: contig of 1697 bp in length  
 \* 105350 105449: gap of unknown length  
 \* 105450 107366: contig of 1917 bp in length  
 \* 107367 107466: gap of unknown length  
 \* 107467 108976: contig of 1510 bp in length  
 \* 108977 109076: gap of unknown length  
 \* 109077 110191: contig of 1115 bp in length  
 \* 110192 110292: contig of 1294 bp in length  
 \* 110292 111586: gap of unknown length  
 \* 111586 112857: contig of 1172 bp in length  
 \* 112857 112957: gap of unknown length  
 \* 112958 114494: contig of 1537 bp in length  
 \* 114495 114594: gap of unknown length  
 \* 114595 116156: contig of 1562 bp in length  
 \* 116157 116256: gap of unknown length  
 \* 116257 117501: contig of 1245 bp in length  
 \* 117501 117601: gap of unknown length  
 \* 117602 118752: gap of unknown length  
 \* 118753 118852: gap of unknown length  
 \* 118853 119859: contig of 1017 bp in length  
 \* 119870 119969: gap of unknown length  
 \* 119970 120986: contig of 1017 bp in length  
 \* 120987 121086: gap of unknown length  
 \* 121087 122100: contig of 1014 bp in length  
 \* 122101 122200: gap of unknown length  
 \* 122201 123600: contig of 1400 bp in length  
 \* 123601 123700: gap of unknown length  
 \* 123701 124857: contig of 1157 bp in length  
 \* 124858 124957: gap of unknown length  
 \* 124958 126055: contig of 1098 bp in length.

## FEATURES

Location/Qualifiers  
 1..126055  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-11F8"

BASE COUNT 35740 a 24872 c 25268 g 35896 t 4279 others  
 ORIGIN

## alignment\_scores:

Quality: 79.50 Length: 45  
 Ratio: 2.650 Gaps: 2  
 Percent Similarity: 66.667 Percent Identity: 44.444

## alignment\_block:

US-09-528-682-3 x AC095291

Align seg 1/1 to: AC095291 from: 1 to: 126055

59 GlsSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGI 75

42133 GAAGAAATAGGGAATGACACCTTCCTAATAGTCCCAATATAA 42182

75 upheLysGlyVal...IleSerAlaAsnValLys.....TyrAspMetA 89

42183 ATACCTCGGTGTGACTCTAGCCCAAGCAAGTGAAGATCTGTATGATA 42232

89 snPheLysLysLeuLeuArgPheMetGluAspasp 100

42233 ACITTAGCTCTGAGAAAGAAATTTGAAGAAGAC 42267

seq\_name: gb\_htg:AC097044

seq\_documentation\_block:

LOCUS AC097044 244530 bp DNA linear HTG 20-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-135H5, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 114 unordered pieces.

ACCESSION AC097044

VERSION AC097044.4 GI:17973524

KEYWORDS HTG: HTGS\_PHASE1.

SOURCE Norway rat.

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus

## REFERENCE

1 (bases 1 to 244530)

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Louissegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S.,  
 Oguh,M., Okwuonu,G., Oraqunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokan,I., Rolfe,M.,  
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshchari,N.,  
 Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Watlington,S., Williams,G., Williamson,A., Wleczylka,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

## TITLE

Direct Submission

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 244530)

## AUTHORS

Worley,K.C.



TITLE  
JOURNAL

Direct Submission  
Submitted (06-OCT-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:17064415.

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GH2S  
Center clone name: CH230-135H5  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329First call to  
findPhrapList

Consensus quality: 187790 bases at least Q40  
Consensus quality: 204453 bases at least Q30  
Consensus quality: 219075 bases at least Q20  
Estimated insert size: 190140; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 1.8x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 114 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 5013: contig of 5013 bp in length  
\* 5014 5113: gap of unknown length  
\* 5114 10621: contig of 5508 bp in length  
\* 10622 10721: gap of unknown length  
\* 10722 16522: contig of 5801 bp in length  
\* 16523 16622: gap of unknown length  
\* 16623 20241: contig of 3619 bp in length  
\* 20242 20341: gap of unknown length  
\* 20342 23852: contig of 5511 bp in length  
\* 23853 25952: gap of unknown length  
\* 25953 29914: contig of 3962 bp in length  
\* 29915 30014: gap of unknown length  
\* 30015 33098: contig of 5084 bp in length  
\* 33099 35198: gap of unknown length  
\* 35199 38096: contig of 2898 bp in length  
\* 38097 38196: gap of unknown length  
\* 38197 41600: contig of 3404 bp in length  
\* 41601 41700: gap of unknown length  
\* 41701 45846: contig of 4146 bp in length  
\* 45847 45946: gap of unknown length  
\* 45947 49106: contig of 3160 bp in length  
\* 49107 49206: gap of unknown length  
\* 49207 52999: contig of 3793 bp in length  
\* 53000 53099: gap of unknown length  
\* 53100 57748: contig of 4649 bp in length  
\* 57749 57848: gap of unknown length  
\* 57849 60767: contig of 2919 bp in length  
\* 60768 60867: gap of unknown length  
\* 60868 63294: contig of 2427 bp in length  
\* 63295 63394: gap of unknown length  
\* 63395 63351: contig of 1957 bp in length  
\* 63352 63451: gap of unknown length  
\* 63452 67346: contig of 1895 bp in length  
\* 67347 67446: gap of unknown length  
\* 67447 69490: contig of 2044 bp in length  
\* 69491 69590: gap of unknown length  
\* 69591 72353: contig of 2763 bp in length  
\* 72354 72453: gap of unknown length  
\* 72454 74557: contig of 2104 bp in length  
\* 74558 74657: gap of unknown length

\* 74658 76789: contig of 2132 bp in length  
\* 76790 76889: gap of unknown length  
\* 76890 79286: contig of 2397 bp in length  
\* 79287 79386: gap of unknown length  
\* 79387 80951: contig of 1465 bp in length  
\* 80952 82392: contig of 1441 bp in length  
\* 82393 82492: gap of unknown length  
\* 82493 84725: contig of 2233 bp in length  
\* 84726 84825: gap of unknown length  
\* 84826 87083: contig of 2258 bp in length  
\* 87084 87183: gap of unknown length  
\* 87184 89485: contig of 2302 bp in length  
\* 89486 89585: gap of unknown length  
\* 89586 92081: contig of 2396 bp in length  
\* 91982 92082: gap of unknown length  
\* 92083 94213: contig of 2132 bp in length  
\* 94214 94313: gap of unknown length  
\* 94314 96608: contig of 2295 bp in length  
\* 96609 96708: gap of unknown length  
\* 96709 100499: contig of 3791 bp in length  
\* 100500 100599: gap of unknown length  
\* 100600 102708: contig of 2109 bp in length  
\* 102709 102808: gap of unknown length  
\* 102809 106610: contig of 3802 bp in length  
\* 106611 106710: gap of unknown length  
\* 106711 108858: contig of 2148 bp in length  
\* 108859 108958: gap of unknown length  
\* 108959 110469: contig of 1511 bp in length  
\* 110470 110569: gap of unknown length  
\* 110570 112253: contig of 1684 bp in length  
\* 112254 112353: gap of unknown length  
\* 112354 115277: contig of 3174 bp in length  
\* 115278 115627: gap of unknown length  
\* 115628 118511: contig of 2884 bp in length  
\* 118512 118611: gap of unknown length  
\* 118612 121134: contig of 2523 bp in length  
\* 121135 121234: gap of unknown length  
\* 121235 123167: contig of 1933 bp in length  
\* 123168 123267: gap of unknown length  
\* 123268 125869: contig of 2402 bp in length  
\* 125870 125769: gap of unknown length  
\* 125770 128562: contig of 2793 bp in length  
\* 128563 128662: gap of unknown length  
\* 128663 130837: contig of 2175 bp in length  
\* 130838 130937: gap of unknown length  
\* 130938 133056: contig of 2119 bp in length  
\* 133057 133156: gap of unknown length  
\* 133157 134203: contig of 1047 bp in length  
\* 134204 134303: gap of unknown length  
\* 134304 136065: contig of 1762 bp in length  
\* 136066 136165: gap of unknown length  
\* 136166 137953: contig of 1788 bp in length  
\* 137954 138053: gap of unknown length  
\* 138054 139875: contig of 1822 bp in length  
\* 139876 139975: gap of unknown length  
\* 139976 142217: contig of 2242 bp in length  
\* 142218 142317: gap of unknown length  
\* 142318 143370: contig of 1253 bp in length  
\* 143371 143670: gap of unknown length  
\* 143671 145446: contig of 1776 bp in length  
\* 145447 145546: gap of unknown length  
\* 145547 147053: contig of 1507 bp in length  
\* 147054 147153: gap of unknown length  
\* 147154 149070: contig of 1917 bp in length  
\* 149071 149170: gap of unknown length  
\* 149171 150389: contig of 1219 bp in length  
\* 150390 150489: gap of unknown length  
\* 150490 152062: contig of 1573 bp in length  
\* 152063 152162: gap of unknown length  
\* 152163 154466: contig of 2304 bp in length  
\* 154467 154566: gap of unknown length



```

alignment_scores:
    Quality: 79.50      Length: 61
    Ratio: 2.338       Gaps: 3
    Percent Similarity: 55.738   Percent Identity: 39.344

alignment_block:
US-09-528-682-3 x AC097044 ..

Align seg 1/1 to: AC097044 from: 1 to: 244530

56 ThrValAsnGlnSerAsnPheProMetSerThr..... 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
12920 ACATTAACATAACAATAAGCTTTCTTACTCAAGAAGATACAGGCT 12969

67 .....PheGluGlnValProAsnAsnLysG 75
||| ||| ||| ||| ||| ||| ||| ||| :
12970 GAGAAAAAATGGGAATGACACCCCTTCACAATAGTCCCAATAATGTAA 13019

75 luPheLysGlyVal...IleSerAlaAsnValLys.....TyrAspMet 88
:||| ||||| :::: ||||| |||||
13020 AATTCTCGGTGTGACTTTAATCAAGCAAGTGAAAGATCTGTATGACAAG 13069

89 AsnPhelYsLysLeuLeuArgPheMetGluasp 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13070 AACCTCAAGTCTCTGAAGAAAGAAATGAAGAT 13102

seq_name: gb_AP003587

seq_documentation_block:
LOCUS AP003587 343550 bp DNA linear BCT 28-NOV-2001
DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 7/19.
ACCESSION AP003587 BA000019
VERSION AP003587.1 GI:17130808
KEYWORDS
SOURCE Nostoc sp. PCC 7120 DNA.
ORGANISM Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
REFERENCE 1 (sites)
AUTHORS Kaneko.T., Nakamura.Y., Wolk.C.P., Kuritz.T., Sasamoto.S.,
Watanabe.A., Iriiguchi.M., Ishikawa.A., Kawashima.K., Kimura,T.,
Kishida,Y., Kohara,M., Matsumoto.M., Matsuno,A., Muraki,A.,
Nakazaki,N., Shindo,S., Sugimoto.M., Takazawa.M., Yamada,M.,
Yasuda.M. and Tabata,S.
Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120
JOURNAL DNA Res 8 (5), 205-213 (2001)
MEDLINE 21595285
PUBMED 11759840
REFERENCE 2 (bases 1 to 343550)
AUTHORS Kaneko.T.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research, Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/cyanobase/,
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
FEATURES
source
location/Qualifiers
1..343550
/organism="Nostoc sp. PCC 7120"
/db_xref="taxon:103690"
/note="synonym:Anabaena sp. PCC7120"
complement(94..411)
/gene="all1719"
complement(94..411)
/gene="all1719"
/note="ORF_ID:all1719
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73418.1"
/db_xref="GI:17130809"
/translation="MRKKSAILEAVHETADLEHAKGLMNQTLIREFHLCPLPIEPL

gene
complement(504..782)
/gene="asl1720"
complement(504..782)
/gene="asl1720"
/note="ORF_ID:asl1720
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73419.1"
/db_xref="GI:17130810"
/translation="MKIYKNRTFDWRARKEGLKNLSICNAVEMAPGLYDADLGGLGF
KKRIAPCKGKGSGGFRLTIATNNEDVGFLFAOKTNAVTLTTHMKKL"
935..1507
/gene="alr1721"
935..1507
/gene="alr1721"
/note="ORF_ID:alr1721"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB73420.1"
/db_xref="GI:17130811"
/translation="MRHKDNKKNQAIQDAIEELITANGFADTSMSKIKAANVSPTAI
YVFENKELLNKIYLFWKISAEILKGVNQLNSVEKAFKMWNYYKYAIKNPRVF
AFTEOFANSPPMDRICKDESLSYFKPMLDLFNRGKEKVKFDIPLEIFTATFEAPLTQ
LIKHFSCALILDEKMLETTYKIAWDATVN"
1569..2504
/gene="alr1722"
1569..2504
/gene="alr1722"
/note="ORF_ID:alr1722
probable oxidoreductase"
/codon_start=1
/transl_table=11
/protein_id="BAB73421.1"
/db_xref="GI:17130812"
/translation="MKNEKWNAENIPPOKGRLAIVTGSSSIGIGVETARVLANKOASVI
TAVRNLDGNKALAKILIQONKDADVKLMELDLANLSVKNPAENFRKNYLRLDLLINN
DFDDLNWPQRYSYAKWAYGDSKLANLYTFELDRKLDNGIDITLVLTASHPGWTATELO
TAGGIVKYLINGIVAQDITMGALPTLRAAIEAGLKGAEYFGPNGFMERMGYPYIKVESN
ELSKDQALAKKLWVWSBKLTVDKFENKKAQSAGK"
complement(2701..3462)
/gene="all1723"
complement(2701..3462)
/gene="all1723"
/note="ORF_ID:all1723
probable aldolase"
/codon_start=1
/transl_table=11
/protein_id="BAB73422.1"
/db_xref="GI:17130813"
/translation="MSVFSRPLPVFNISIIEERLHRKORLAAAARFLPKFGFSEGIAG
HITARLPDTDFHWPNPLGYFCHIRVSDLVNRDGDVVKGDALINQAFAIHSHOI
EARPDVAHAHASTYCKAWSSLGLRDPLTDQACAFVEDHALDDYTGVVLTSGEK
BLAETLGNNAKALLILRNHGILTVGHFTVDEAAFVSIERSCOAQLLAABAOPKFIIGE
TARLTHSQVGEPKAGWFSPQLDYDVAREEFPELD"
complement(3764..5167)
/gene="all1724"
complement(3764..5167)
/gene="all1724"
/note="ORF_ID:all1724
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73423.1"
/db_xref="GI:17130814"
/translation="MSSTSQVGPGLDPHTTELIOGLSOLITASLLLVFLVALGTA
IALMSFALRHNSOOTIFIGEWARIYAQLRHLQHLALVILLVGGFFLSSTLSNRHY
HWDAQIOVVQSVSGDKLEQIAPQVRYVTQNYGVNGKIVKVNDTPQQVSRLLS
LAGSQIAQVKKBSQDQGSRAVADYATDYKVYNRLKIDINNEFFEIQPPNGYSLLS
```

EPLQIKETRESSQVSQAVFARILNLISTSTVQKWEIGOKRFGSGASLKLLHLVLVKNRGLN

VLY"

gene complement(504. .782)

CDS /gene="asl1720"

complement(504. .782)

/gene="asl1720"

/note="ORF\_ID:asl1720

hypothetical protein"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB73419.1"

/db\_xref="GI:17130810"

/translation="MKIYKNRTDFWRKKEGLKNLSCNAVEMAPGCLYDADLGGLGF

KKRIAKPGKGSGGFRTLATINNEDVGFLFLAQKTNAVTLTTHMKKL"

935. .1507

/gene="alr1721"

CDS 935. .1507

/gene="alr1721"

/note="ORF\_ID:alr1721"

/codon\_start=1

/transl\_table=11

/product="transcriptional regulator"

/protein\_id="BAB73420.1"

/db\_xref="GI:17130811"

/translation="MRHKDNKKNOAICDAEILTANGFADTSMKSITAKAAANVPATI

YVFENKEKLNIYLFVKQSISAEILKGVNQNLSEKAFKMWNYYKIAYIKNPVR

AFTQFANSMPWDRICKDESLSYFKPMLDLFNRGKEKVEKDIPLEIFTAFTAFLPTL

LKHFSGALILDEKMLETTYKIAWDATN"

1569. .2504

/gene="alr1722"

CDS 1569. .2504

/gene="alr1722"

/note="ORF\_ID:alr1722

probable oxidoreductase"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB73421.1"

/db\_xref="GI:17130812"

/translation="MKNEKWNAENIPPOKGRLAIVTGSSSIGIVETARVLANKOASVI

TAVNLDGNKALAKILQQNKDADVCLKMELDLNLASVKNAFENPKNYLRLLDLINN

AFDDLPPIYSKYTDGELOFGTNHGLHFALTGQLLEFLISTEGSRIVNVSSGAHNMGKI

DFDDLNWQRSAWKAYGDNSKLANLYTFELDRLKNDNGIDITLVTAASHPGWTATEIQ

TAGGIVKYLINGIVAODITMGALPTLRAAIEAGLKGAEYFGPNGFMERMGPYPKIVESN

ELSKDQALAKLWVWSBKLTVDKFENKKAQSAGK"

complement(2701. .3462)

/gene="all1723"

CDS complement(2701. .3462)

/gene="all1723"

/note="ORF\_ID:all1723

probable aldolase"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB73422.1"

/db\_xref="GI:17130813"

/translation="MSVFSRPLPVFNISIIEERLHRKORLAAAARFLRKFPGFSEGIAH

HITARLDPETHFWNWPLGKVGHIRVSDLVVNRRDGVVKGDALINQAFAIHSHOI

EARDPVAAAHASTYCKAWSSLGLRLDPLTDACAFVEDHALLDYTVGVVLETSGEK

BLAETLGNNAKALLILRNHGILTVGHTVDEAAFVSIERSCOAQLLAABAOPKEIGH

TARLTHSQVGEPKAGWFSFQPLDYRVAREPELFD"

complement(3764. .5167)

/gene="all1724"

CDS complement(3764. .5167)

/gene="all1724"

/note="ORF\_ID:all1724

unknown protein"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB73423.1"

/db\_xref="GI:17130814"

/translation="MSSTSQVPGPLDPTHTELIQIGLSOLITASLLLVFLVIALGTA

IALMSFALRHNSOOTIFIGEWARIYAQILRHQHLALVILLVLGGFFLSSTLSNRHY

HWDAQIOVVQSVGSQDKLEQIAPQVRYVTQGYPTNTQVNGKIVKYNVDTPQQVSRLLS

LAGSQIAQVVKLBQSDQGRSAVRADYTADYKVVNRLKIDINNEFFEQPPNGYSLLSN



```
YKVERDTRLEPVPNGVSYFAFRLEPGOETSRVYTOAQGGPRWVINSAGOLLNFRLL
TTIANFGADPAGSIVPKERKADGRSTQYTWIFEDNTSVKKNPKGCVFTNTNPIROTGII
PRILLALLPALEFLWILLIXLSLPISLINVAIASSIYFACILTLILYLSRFPDRAQATW
ISLVLLITWGGSGSTRASLAALICTTAGAVVPLGLVPLYPYSGLTLSVAGLLSVIWL
VRHYNWNIWOPEDSKS"
complement(5492. .5902)
/gene="all1725"
CDS
complement(5492. .5902)
/gene="all1725"
unknown protein"
/note="ORF_ID:all1725"
/codon_start=1
/transl_table=11
/protein_id="BAB73424.1"
/db_xref="GI:17130815"
/translation="MFLKPSGLMVIASLTNGIAITYSGRSATAEAKSNFYLOVGA
DQGFPTALDIASINGNTLVQQHGGIARRLHAACSGRLFTKFKPSLYAANGKLT
SERTEQIEFPKPTPDAASEIVEICRVANNHQQK"
6321. .6842
/gene="alr1726"
CDS
6321. .6842
/gene="alr1726"
/note="ORF_ID:alr1726"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73425.1"
/db_xref="GI:17130816"
/translation="MKSQHLQATTLAGEVSQOYLKEEFPALNAIAELQFIDMCPDARE
VRKALVLYQGYLYDEIQILDSRGISITGKWQAYEETGIDGLRLNKGKSYLNA
KOLDEVLSWLQTKDCWELGELEYOLAFEYDVVYESQSYVYDLFSEAGISWKKTKVNP
6319. .7431
/gene="alr1727"
CDS
6319. .7431
/gene="alr1727"
/note="ORF_ID:alr1727"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB73426.1"
/db_xref="GI:17130817"
/translation="MWGDLGYSVMGKSDQELAVPVNDRKOTYGAVDYLLGLSELVK
AYDAGNSKNTNLYLESLANSPODLRLIFWDGASVHRSKEIRGLFSDVNOQLPTBQWK
IHCYREAPNCVPQNPIDILWLOAKTWRRFCALPSPSHLKWIFENFIRHTTFDEDL
QMYGTYSKIK"
8064. .9401
/gene="alr1728"
CDS
8064. .9401
/gene="alr1728"
/note="ORF_ID:alr1728"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73427.1"
/db_xref="GI:17130818"
/translation="WTKGNARHNAFLRLVSGNGAAGFSGESRYLSLTSKEVWIGRDPSCQ
VYLDAMTRWVSRHRAVAVAPQPTIITNQVAPLPSAKHPKPDVSFTQLFPIITGKD
DRISIGADPQFLFEYAVAPQPTIITNQVAPLPSAKHPKPDVSFTQLFPIITGKD
LTRKALYVPGITLVVFMFATVGRPOANQVIVATYALAAVFIYRLCGKPKPWW
LVGAALSTLILSPLLFIPIREVLPGSVPIIDQPVSLTELFVYFFGAGLME
LKALPIGLVYLIALGLPAPWRVGVWEPLOGILLGSTASAVGFTLLETGLQYVPAAS
LQATVEGLQVLIARLGLPAGHAYSGYLGFFGLAALKPRHAKQIILAVGLYSAL
HALMNTAGHSNNLLIVVGVLSYAFPLMAAILKARALSPTRSONFAITRFLDPKS"
complement(9666. .10310)
/gene="all1729"
CDS
complement(9666. .10310)
/gene="all1729"
/note="ORF_ID:all1729"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB73428.1"
```

```
/db_xref="GI:17130819"
/translation="WTKHTNGSDACHSVVSQWELDLLAALLEPEDAAYPNWPSDDDESE
TYFELEKRFQDVSDSEFTTQAFYFDKLDHLSGVTTPSSYINTTISVYVDHLQEN
LHTAFATIPQNLNIAIATKAAEIFASQOGLSEQVLCVSVLPTWEADDDLLIFARP
AYSRRSSQGLTSTIVRNVNQWSELSALSEVEQAKLSVAIAYAIROLNDSPPSE"
complement(10795. .12372)
/gene="all1730"
CDS
complement(10795. .12372)
/gene="all1730"
/note="ORF_ID:all1730"
```

```
alignment_scores:
  Quality: 79.50      Length: 118
  Ratio: 1.223      Gaps: 6
  Percent Similarity: 55.085      Percent Identity: 24.576

alignment_block:
US-09-528-682-3 x AP003587 ..
Align seg 1/1 to: AP003587 from: 1 to: 343550

1 AspPhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAs 17
  ::::||||| ::::: ||| ||:::
324695 GAGATTTTATTAAACGATGGCGCAGCGTACCACAA..... 324730

17 nThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAsp. 33
  ||||| ::::: ||||| ||||| :::::
324731 .....GTGACTCTTTTGGGTGGTAGTTTCCCGTACCGGTAAAGTG 324770

34 .....GlyArg...TyrSerAsnPheAlaLeuIleSerAlaGluGly... 46
  ||||| ::::: ||||| :::::
324771 ATACAGGTAGAGTCTATACACACCACCATTCATCATCAGTCCCGGTGAA 324820

47 .....GlyMetGlnAs 50
  :::: ||
324821 GAACCTTGCTCGCTACAATAAGTACACCTTTTGTGATGTCATGTCCTGA 324870

50 pGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrP 67
  ||||| ::::: ||||| :::::
324871 TGGTAATACCTAC.....CGCACTAGTACTGTTGTCGCCG 324908

67 heGluGlnValPro.....AsnAsnLysGluPheLysGlyValIleSer 81
  :::: ||||| ::::: ||||| :::::
324909 GTCAGCAATTACCCCGTCCATTTTTCGGAATATCTAGCAATATCGGC 324958

82 AlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPheMetG1 98
  :::: ||||| ::::: ||||| :::::
324959 GTTCTATTGCTACGATGCCGCTTCCCGAAGTCTACCGCATCTCTC 325008

98 uAsp 99
  :|||
325009 AGAC 325012
```

```
seq_name: gb_ov:AB008162
seq_documentation_block:
LOCUS AB008162 2178 bp mRNA linear VRT 11-JUN-1998
DEFINITION Xenopus laevis N-CAM 1 mRNA for neural cell adhesion molecule,
complete cds.
ACCESSION AB008162
VERSION AB008162.1 GI:3116226
KEYWORDS neural cell adhesion molecule; N-CAM 1.
SOURCE Xenopus laevis adult heart cDNA to mRNA.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2178)
Kudo,M. and Shiohawa,K.
Direct Submission
Submitted (16-OCT-1997) Mariko Kudo, Graduate School of Science,
University of Tokyo, Biological Sciences, Hongo 7-3-1, Bunkyo-ku,
Tokyo 113-0033, Japan (Tel:03-3812-2111(ex.4433), Fax:03-3816-1965)
```



```

REFERENCE 2 (bases 1 to 2178)
AUTHORS Kudo,M., Takayama,E., Tadakuma,T. and Shikawa,K.
TITLE Molecular cloning of ssd-form neutral cell adhesion molecules
(N-Cams) as the major form in Xenopus heart
JOURNAL Biochem. Res. Commun. 245 (1), 127-132 (1998)
MEDLINE 98204770
FEATURES             Location/Qualifiers
     source          1..2178
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /tissue_type="heart"
     gene            1..2178
                     /gene="N-CAM 1"
     CDS              1..2178
                     /gene="N-CAM 1"
                     /codon_start=1
                     /product="neural cell adhesion molecule"
                     /protein_id="BAA25931.1"
                     /db_xref="GI:3116227"
                     /translation="MLHFKDLIWTLYFGTAVALEVNIYPDQGEISLGSFKFLCQVS
                     GEATDISWYSPTEKILVTQOOISVVRSDDYTSLTIYNASSQDAGIKCVASNEAEGE
                     SEGTNLKIYQKLTFRNAPTQEFKEGDAVICDVSSIPSITWRHKGKDVIFAKD
                     VRFVLANNYLQIRGIKKTDGTYRCEGRILARGEIKDIQIVNVPTTIOARLVR
                     NATAKMAESVLSCDADGFPDPEISWLKKEPIEDGEERKISFNEQDSEMTIHVVKDRV
                     EAYSCIANNOAGEAEATILLKYAKPKITYVENKTAVELDEITLTCEASGDPISIT
                     WRTAVRNISSEATLLDGHIVVKEHIRMALTLKDIOYTDAGEVFCIASNPICGVDMQAM
                     YFEVOYAPKIRGPVYVWEGNPVNTICEVFAHRAAVTFRDGGQLLPSNPSNLIKIY
                     SGTSSLEVPDSENFNGYNTAINTIGHEFSEFIVQADTPSPAIRKVEPYST
                     VMVTFDEPSTGVPILKYKAERWVIGHEKWHYKYDAKEVNAESITVMGLKPTSY
                     MVKLSAMNGKLGSDTPSQEFTTQPVYISKPGCEAPKLVGLHLSDEGNSIRKVDILKQ
                     DGGSPRIYLVNRYALNALNAPKMPRVPSNHHVLMKALEWNVYEVIVVAENQOQK
                     SKPALLSFRTAKPTATTANVGSSASSNVVPLLVLCLVIAIVPC"
BASE COUNT      685 a 453 c 487 g 553 t
ORIGIN
alignment_scores:
    Quality: 79.00      Length: 108
    Ratio: 1.339      Gaps: 5
    Percent Similarity: 54.630      Percent Identity: 28.704

alignment_block:
US-09-528-682-3 x AB008162 ..
Align seg 1/1 to: AB008162 from: 1 to: 2178
18 ThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspG1 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 ACTTCACACTCACCATCTACAATGCCAGCAGCAAGATGCT.....GG 266
34 yArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspG 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 CATCTATAATGTGTAGCTTCCACGAGGACAGAGGAGAACTCGAA.... 312
51 lyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPhe 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 .....GGCACTGCTCAATCTTAGATTATCAGAAATGACCTTT 351
68 GluGlnValProAsnAsnLysGluPheLys.....GlyValI1 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 AAAAAATGACCCACCCCTCAGGAGTTTAAAGAGGAGAGATGCGATCAT 401
80 eSerAlaAsnValLys..... 85
402 TATTTGTGATGTCCTCACTCCATTCCTTCAATTATCATCTGGGACATA 451
86 ....TyrAspMetAsnPheLysLysLeuLeuArgPhe.....MetGlu 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 AAGGCAAGATGTTATTTTCAAAAAAGATGTACGGTTTGTGTCTTGCC 501
99 AspAspPheIleGlyValHisGly 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

502 AACAAATTACCTCCAAATCAGGGGA 525
seq_name: gb_ov:XELNCAM
seq_documentation_block:
LOCUS XELNCAM
DEFINITION X.laevis neural cell adhesion molecule (NCAM) mRNA, complete cds.
ACCESSION M25696
VERSION M25696.1 GI:214609
KEYWORDS neural cell adhesion molecule; transmembrane protein.
SOURCE X.laevis neurosula, cDNA to mRNA, clone N1.
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 3720)
AUTHORS Krieg,P.A., Sakaguchi,D.S. and Kintner,C.R.
TITLE Primary structure and developmental expression of a large cytoplasmic domain form of Xenopus laevis neural cell adhesion molecule (NCAM)
JOURNAL Nucleic Acids Res. 17, 10321-10335 (1989)
MEDLINE 90098871
COMMENT Draft entry and computer-readable sequence [1] kindly submitted by P.A.Krieg, 23-JUN-1989.
FEATURES             Location/Qualifiers
     source          1..3720
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
     sig_peptide      23..81
                     /note="neural cell adhesion molecule signal peptide"
     CDS              25..3291
                     /note="neural cell adhesion molecule precursor"
                     /codon_start=1
                     /protein_id="AAA49909.1"
                     /db_xref="GI:214610"
                     /translation="MLHFKDLIWTLYFGTAVALEVNIYPDQGEISLGSFKFLCQVS
                     GEATDISWYSPTEKILVTQOOISVVRSDDYTSLTIYNASSQDAGIKCVASNEAEGE
                     SEGTNLKIYQKLTFRNAPTQEFKEGDAVICDVSSIPSITWRHKGKDVIFAKD
                     VRFVLANNYLQIRGIKKTDGTYRCEGRILARGEIKDIQIVNVPTTIOARLVR
                     NATAKMAESVLSCDADGFPDPEISWLKKEPIEDGEERKISFNEQDSEMTIHVVKDRV
                     EAYSCIANNOAGEAEATILLKYAKPKITYVENKTAVELDEITLTCEASGDPISIT
                     WRTAVRNISSEATLLDGHIVVKEHIRMALTLKDIOYTDAGEVFCIASNPICGVDMQAM
                     YFEVOYAPKIRGPVYVWEGNPVNTICEVFAHRAAVTFRDGGQLLPSNPSNLIKIY
                     SGTSSLEVPDSENFNGYNTAINTIGHEFSEFIVQADTPSPAIRKVEPYST
                     VMVTFDEPSTGVPILKYKAERWVIGHEKWHYKYDAKEVNAESITVMGLKPTSY
                     MVKLSAMNGKLGSDTPSQEFTTQPVYISKPGCEAPKLVGLHLSDEGNSIRKVDILKQ
                     DGGSPRIYLVNRYALNALNAPKMPRVPSNHHVLMKALEWNVYEVIVVAENQOQKSKPALL
                     SFRTAKPTATTATASAGTGLGTGAIVGLIVFVILLVVDVDTCTFLNCKGSKLALIA
                     VNFCCGAGPAGKAGKDIIEGKAASFKESEKPIVEVRETEETPNHDGNSIPEPNTTP
                     LITEPEHPAAVEDMLPSVTVTNTSDITITETATQNSDTSETTTLTSTAPPPTTAPD
                     SNTIOSIQTAPSKAEAPTTSSPPTSSPKVAPLVDLSDPTNNPSKVVAQAQGLNPS
                     AATSAEPTVTIIKPVTVPPNAASPPPTPEPKVQKQSGTSPKEKEAOPSTVKNP
                     TEATKDESASLNTKPLQDEDFQIDGGTFTKPEIOLAKDVFALGTATPTAVASCKAS
                     ELVSSTADTSPVLDKSAKTEKQVEEKSPEDVAGTFAEVKTVFNEATQTNANESKA
                     "
     mat_peptide      82..3288
                     /product="neural cell adhesion molecule"
BASE COUNT      1158 a 851 c 801 g 910 t
ORIGIN 1 bp upstream of EcoRI site.
alignment_scores:
    Quality: 79.00      Length: 108
    Ratio: 1.339      Gaps: 5
    Percent Similarity: 54.630      Percent Identity: 28.704

alignment_block:
US-09-528-682-3 x XELNCAM ..
Align seg 1/1 to: XELNCAM from: 1 to: 3720
18 ThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspG1 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```











Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S.,  
 Oguh, M., Okwuonu, G., Otagunye, N., Oviedo, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,  
 Qullis, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,  
 Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,  
 Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Gibbs, R.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Direct Submission  
 2 (bases 1 to 182686)  
 Worley, K.C.  
 Direct Submission  
 Submitted (23-Oct-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:17062348.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GGBB  
 Center clone name: CH230-167C14  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 158745 bases at least Q40  
 Consensus quality: 165166 bases at least Q30  
 Consensus quality: 170378 bases at least Q20  
 Estimated insert size: 160185; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 2.3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 64 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 10034: contig of 10034 bp in length  
 \* 10035 10134: gap of unknown length  
 \* 10135 16324: contig of 6190 bp in length  
 \* 16325 16424: gap of unknown length  
 \* 16425 22385: contig of 5961 bp in length  
 \* 22386 22485: gap of unknown length  
 \* 22486 29367: contig of 6882 bp in length  
 \* 29368 29467: gap of unknown length  
 \* 29468 34112: contig of 4645 bp in length  
 \* 34113 34212: gap of unknown length  
 \* 34213 39026: contig of 4814 bp in length  
 \* 39027 39126: gap of unknown length  
 \* 39127 44825: contig of 5699 bp in length  
 \* 44826 44925: gap of unknown length  
 \* 44926 49550: contig of 4625 bp in length  
 \* 49551 49650: gap of unknown length  
 \* 49651 54932: contig of 5282 bp in length  
 \* 54933 55032: gap of unknown length  
 \* 55033 59461: contig of 4129 bp in length  
 \* 59462 65116: contig of 5855 bp in length  
 \* 65117 65216: gap of unknown length  
 \*  
 \* 65217 70187: contig of 4970 bp in length  
 \* 70187 70286: gap of unknown length  
 \* 70287 74102: contig of 3816 bp in length  
 \* 74103 74202: gap of unknown length  
 \* 74203 77264: contig of 3062 bp in length  
 \* 77265 81515: contig of 4151 bp in length  
 \* 81516 84785: gap of unknown length  
 \* 84786 84886: contig of 3170 bp in length  
 \* 84887 87347: gap of unknown length  
 \* 87348 91366: contig of 2461 bp in length  
 \* 91367 91466: gap of unknown length  
 \* 91467 95032: contig of 3565 bp in length  
 \* 95033 95132: gap of unknown length  
 \* 95133 97787: contig of 2656 bp in length  
 \* 97788 100330: contig of 2443 bp in length  
 \* 100331 104094: contig of 3664 bp in length  
 \* 104095 104194: gap of unknown length  
 \* 104195 106403: contig of 2209 bp in length  
 \* 106404 106503: gap of unknown length  
 \* 106504 109021: contig of 2518 bp in length  
 \* 109022 109121: gap of unknown length  
 \* 109122 110855: contig of 1744 bp in length  
 \* 110856 110966: gap of unknown length  
 \* 110967 113374: contig of 2408 bp in length  
 \* 113375 113474: gap of unknown length  
 \* 113475 115960: contig of 2487 bp in length  
 \* 115961 116061: gap of unknown length  
 \* 116062 118527: contig of 2467 bp in length  
 \* 118528 120200: contig of 1573 bp in length  
 \* 120201 122173: contig of 1873 bp in length  
 \* 122174 122374: gap of unknown length  
 \* 122375 124447: contig of 2074 bp in length  
 \* 124448 126947: gap of unknown length  
 \* 126948 127047: gap of unknown length  
 \* 127048 129787: contig of 2740 bp in length  
 \* 129788 131648: contig of 1761 bp in length  
 \* 131649 131748: gap of unknown length  
 \* 131749 133414: contig of 1666 bp in length  
 \* 133415 135315: contig of 1801 bp in length  
 \* 135316 138035: gap of unknown length  
 \* 138036 138135: contig of 2620 bp in length  
 \* 138136 139491: gap of unknown length  
 \* 139492 141470: contig of 1879 bp in length  
 \* 141471 141570: gap of unknown length  
 \* 141571 143465: contig of 1895 bp in length  
 \* 143466 143565: gap of unknown length  
 \* 143566 145265: contig of 1600 bp in length  
 \* 145266 147498: contig of 2232 bp in length  
 \* 147499 147597: gap of unknown length  
 \* 147598 149559: contig of 2162 bp in length  
 \* 149560 152108: contig of 2249 bp in length  
 \* 152109 152208: gap of unknown length  
 \* 152209 154101: contig of 1893 bp in length  
 \* 154102 155722: contig of 1521 bp in length  
 \* 155723 157142: contig of 1320 bp in length  
 \* 157143 158824: contig of 1582 bp in length  
 \* 158825 159242: gap of unknown length  
 \* 159243 159824: contig of 1582 bp in length  
 \* 159825 160000: gap of unknown length



\* 158825 158924: gap of unknown length  
 \* 158925 160188: contig of 1264 bp in length  
 \* 160189 160288: gap of unknown length  
 \* 160289 162721: contig of 2433 bp in length  
 \* 162722 162821: gap of unknown length  
 \* 162822 164326: contig of 1405 bp in length  
 \* 164327 164327: gap of unknown length  
 \* 164327 165784: contig of 1458 bp in length  
 \* 165785 165884: gap of unknown length  
 \* 165885 167562: contig of 1678 bp in length  
 \* 167563 167662: gap of unknown length  
 \* 167663 168684: contig of 1022 bp in length  
 \* 168685 168784: gap of unknown length  
 \* 168785 170449: contig of 1665 bp in length  
 \* 170450 170549: gap of unknown length  
 \* 170550 171949: contig of 1400 bp in length  
 \* 171950 172049: gap of unknown length

alignment\_scores:  
 Quality: 78.50 Length: 41  
 Ratio: 2.804 Gaps: 1  
 Percent Similarity: 68.293 Percent Identity: 46.341

## alignment\_block:

US-09-528-682-3 x AC098098/rev ..

Align seg 1/1 to reverse of: AC098098 from: 1 to: 182686

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyValII 80

||||| ||| ||||||||| |||||

115184 ATGACACACTTCATAATAGTCCAAATAATAATAACCTTGGTGTGAC 115135

80 e.....SerAlaasnValLysTyrAspMetAsnPhelLysLeuL 94

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

115134 TTTAACCAAGCAGCGAAAGATGCTGTATGATGAACTCAAGTCTCTGA 115085

94 euArgPheMetGluAspPhe 101

::: ||||| |||||

115084 AGAAGAAATTGAAGACACTTC 115062

seq\_name: gb\_htg.AC094978

seq\_documentation\_block:

LOCUS AC094978 121958 bp DNA linear HTG 20-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-619, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
 68 unordered pieces.

ACCESSION AC094978

VERSION AC094978.2 GI:117941178

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
 Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,J., Newton,N.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,  
 Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,T., Rolfe,M.,  
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,  
 Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Williamson,C.,  
 Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G. and Gibbs,R.

## Direct Submission

Unpublished

2 (bases 1 to 121958)

Worley,K.C.

## Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15624815.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBVQ

Center clone name: CH230-619

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 95065 bases at least Q40

Consensus quality: 103860 bases at least Q30

Consensus quality: 109274 bases at least Q20

Estimated insert size: 67448; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 0.9x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 68 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 3355: contig of 3355 bp in length

\* 3356: gap of unknown length

\* 3456: contig of 2617 bp in length

\* 6073: gap of unknown length

\* 6172: contig of 1446 bp in length

\* 6173: gap of unknown length

\* 7619: contig of 2217 bp in length

\* 7719: gap of unknown length

\* 9936: contig of 2029 bp in length

\* 10036: contig of 2029 bp in length

\* 12065: gap of unknown length

\* 12164: contig of 2306 bp in length

\* 12165: gap of unknown length

\* 14471: contig of 2429 bp in length

\* 14571: gap of unknown length

\* 17000: contig of 3107 bp in length

\* 20206: contig of 3107 bp in length







Huber, R., Feldman, R.A., Short, J.M., Olson, G.J. and Swanson, R.V.  
Direct Submission  
Submitted (25-JUL-1997) Diversa Corporation, Genomics, San Diego,  
CA 92121

Putative indicates no similarity to known proteins

Hypothetical indicates similarity to a protein of unknown function.

Location/Qualifiers

FEATURES

```
source
1..13097
/organism="Aquifex aeolicus"
/strain="VF5"
/db_xref="taxon:63363"
163..966
/gene="kdsA"
/notes="aq_085"
163..966
/gene="kdsA"
/codon_start=1
/transl_table=11
/product="3'-deoxy-d-manno-octulosonic acid 8-phosphate
synthase"
/protein_id="AAC06457.1"
/db_xref="GI:2982835"
/translation="MEKFLVIAAGPCAIASESELLLKVGEIKRLSEKFEVEPVFKSSF
DRANKSSHSFGHLEYGKALRVKVEEGLKITTDIHESWQAEPPVAEADIIPPA
FLCRQDILLAAKTRAVNVKGGQFLAPWDTKNVVEKLFGGAKEIYLTERGTTFGY
NNLVDFRSLPTMKQWAKVIYDATHSVQLPGLGDKSGMGRFIFPLIRAAVAGCDG
VFMETHPEPEKALSDASTQLSLEQIATEILREIVASKYIETIPVK"
963..1625
/gene="modC"
/notes="aq_086"
963..1625
/gene="modC"
/codon_start=1
/transl_table=11
/product="Molybdenum transport system permease"
/protein_id="AAC06461.1"
/db_xref="GI:2982839"
/translation="MKEALFLSKLSFFTTLLILLISLVLSPLAFKFKGKSLVVAL
VTLPLVPTVLGVLILLFSPSPFGKLVQSLIKTLTFLSPFGLVAVSVIYSLPLAV
LPLTSAMEKIPKELIYTSVFCYGVKVEFFRVLPLSLNGLITSVMSPANTMGEGV
VLMVGNIGETQTLISFIYDAVQALDYKTAHRASVLLVGVSLLVNFVILKVRGRV
DV"
1748..2089
/gene="aq_087"
1748..2089
/gene="aq_087"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06465.1"
/db_xref="GI:2982843"
/translation="WATLSKLKELMSIPGAVAGFEESDDGRLLAYVYGDIDEKSAETAA
MCAANKMGMNQAKQWSAYTKDGFYPVIGFAVAGGKYAACIMNGVGVFVELDRADF
DKIFETLSKFI"
2104..2439
/gene="aq_088"
2104..2439
/gene="aq_088"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06466.1"
/db_xref="GI:2982844"
/translation="WANLDKLLQIKGVFAAGFEESDDGKLVAYKGDITEEAAAEAMC
AANNAKMQAGCYTAFSGKEWTPLVGNALTPKYSVCVGVNGVFNDEVSFNEVF
KALREAGK"
2467..2871
/gene="aq_090"
2467..2871
/gene="aq_090"
/codon_start=1
/transl_table=11
/product="putative protein"
```

```
gene
CDS
/protein_id="AAC06467.1"
/db_xref="GI:2982845"
/translation="MEDRWAKYFLSEDWIKAYAEWKNKNEKIMNDLKNFSASIKYVVE
GKENDAVELIIEKXAVSAGRADPNKRYDFELWASLENWKKLATGEMGPRAMLTKRL
RFKSGMITAMKTMGAEEESLKMGRVPTVMDV"
2864..3760
/gene="aq_091"
2864..3760
/gene="aq_091"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06468.1"
/db_xref="GI:2982846"
/translation="MFKFINKVDKFTFDEIFDGLVIVIDKNRRIVPANKKAKEILGEVK
EGESCRGLFSICEHCPMELVEDSEQGVQYDVLTFRKKNHVCLSNSPLSEGEFTGVLE
FRDVSKEVIHMEEVKROKFEVVDLSIVAEVLVDDEKGNVININGNASKVCLSRFE
SITGNLKEITSLSLEDLPPEGERADIYETPCGQKASVLLSKLKQKQGWLSLYIV
PEIFAGEGNKLIKIVTKSPKFLVKLEKVKTISEMNVNVLITGETGTGKSLARYIHYLS
PRDKPFVKINCGAIPENLEA"
3824..4456
/gene="hth"
/notes="aq_093"
3824..4456
/gene="hth"
/codon_start=1
/transl_table=11
/product="transcriptional regulator (H-T-H)"
/protein_id="AAC06459.1"
/db_xref="GI:2982837"
/translation="MANGGTLFLDEIGELPLHLQVKKLGLIQDKEFERIGDLKPRKAD
VRIIAATNKNLKLVEEGKREDLYRLNVNVELPPLRERKEDIPHLYSVLPKFSK
IHGKVKGVSPAMKLLLEYDYPGNVRELENAIEHAVVSVSTSLIGPDPLPEEIRIGS
KYKKFRENEELEKIKALERAGGNKSLAALLGHRTILWRKLEKEYLA"
complement(4453..6858)
/gene="nrda"
/notes="aq_094"
complement(4453..6858)
/gene="nrda"
/codon_start=1
/transl_table=11
/product="ribonucleotide reductase alpha chain"
/protein_id="AAC06460.1"
/db_xref="GI:2982838"
/translation="MTMVYIKRGRKEKLDINKIRIAIKFACGLNVDPLEADAQOI
QFRDGIITKEIQOLLIKTAAEKVSAREPDWTYARLLLDLYDKVAHLRGYSLRDDL
GKIKYNNKNTFSYFKEIVEYIGYEYLLNSENEDPNKLANIKPERDLFTYTGIK
ILYDRLYRDEEGRVIELPQEMYMLIAMTLAVPEKPERLKWAKFYDVLSEHKVTVA
TPTLMNARPTQLSSCFVLTVDDDLDFIDFNKKAGMISKFAGGLGYLGLKIRATGA
PIRKFAGASSGVPVVKLINDTMTYVDQGLMKRKSASITLDIWHKDIIDLFLEKVTNVG
DERKKAHDHPAVSIPDLFMKRLKNREDWTLDIPYARQYITRKYDCKYKEVKPLPG
SHYVYIKEDGTODILEPKGLEDFEYEEFKWYLELENLPSYAKKYNSEPLMKRLL
TVAFTEGPPYIFRDEARNKNPKHTGMYVYNSNLCHEIVQTMSPSKHEKVPILDPTGE
ITYKKEAGDLPVCNIGSVNLGKVHTEETKEVPLLLVRMLDNVIMENFYAIEAEYTN
KRYAIGIGVSNYHYCLVNGIKWSEEHKFAADKLFIAYFALKGLSELAKEGRGY
KLFDGNSNKGILLFGRSVEEIEENRQNNLPMRELAEIKKYIRNAYLIALMPTG
STSLILGATPSIDPIFAFYKEENMSGILPOVPPVEDRYWHYKTYTIDHEWTIRAA
AVRQKWIDQASLNLFLVDPQNDGPRLSRLYELAWELGLKTYIYILRSAMDIECEA
CSV"
complement(6944..8167)
/gene="aq_097"
complement(6944..8167)
/gene="aq_097"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06469.1"
/db_xref="GI:2982847"
/translation="WEILNRLTLTDYIVFTSFILYLLFFLSPFLERKRLWTVAVTPLA
SLICSGLYSAPLLYVYKVAIILGMAGIVILYIGAIYNIKAEPILYGSSESK
LKALLDVERFSNLAFAYMISIAFYRLLSFVSFGFFERNVEYRLLTGILLFI
GISGPIRLHLEFMEKAVAGLKUSIIFSFTALLYIKNPSFAGEVKPFDLHLEV
LGILLIVQGFETSKYLGEYTPPEERVKSMKLAQWISGFIYVSFMTLITSFVFNHPK
```







-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hqsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hqsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 61 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1
* 3232: contig of 3232 bp in length
* 3233: gap of unknown length
* 3333: contig of 2369 bp in length
* 5701: gap of unknown length
* 5702: gap of unknown length
* 5802: gap of unknown length
* 8171: contig of 2370 bp in length
* 8271: gap of unknown length
* 8272: contig of 1977 bp in length
* 10248: contig of 1977 bp in length
* 10348: gap of unknown length
* 10349: contig of 1489 bp in length
* 11838: gap of unknown length
* 11838: contig of 1270 bp in length
* 13207: gap of unknown length
* 13208: gap of unknown length
* 13308: contig of 1442 bp in length
* 14750: gap of unknown length
* 14849: contig of 1645 bp in length
* 14850: gap of unknown length
* 16495: gap of unknown length
* 16595: contig of 2172 bp in length
* 18766: gap of unknown length
* 18767: gap of unknown length
* 18866: gap of unknown length
* 20111: contig of 1245 bp in length
* 20112: gap of unknown length
* 20121: contig of 1388 bp in length
* 21599: gap of unknown length
* 21600: gap of unknown length
* 21600: contig of 1749 bp in length
* 23448: gap of unknown length
* 23449: gap of unknown length
* 23549: contig of 1711 bp in length
* 25259: gap of unknown length
* 25359: gap of unknown length
* 25360: contig of 1090 bp in length
* 26449: gap of unknown length
* 26450: gap of unknown length
* 26550: contig of 1341 bp in length
* 27890: gap of unknown length
* 27891: gap of unknown length
* 29383: contig of 1393 bp in length
* 29384: gap of unknown length
* 29484: contig of 1893 bp in length
* 31376: gap of unknown length
* 31377: contig of 1971 bp in length
* 31477: gap of unknown length
* 33447: gap of unknown length
* 33448: gap of unknown length
* 34632: contig of 1085 bp in length
* 34732: gap of unknown length
* 34733: contig of 1273 bp in length
* 36005: gap of unknown length
* 36006: gap of unknown length
* 36106: contig of 1157 bp in length
* 37262: gap of unknown length
* 37263: contig of 1195 bp in length
* 38557: gap of unknown length
* 38558: gap of unknown length
* 40292: contig of 1635 bp in length
* 40392: gap of unknown length
* 40393: contig of 1115 bp in length
* 41507: gap of unknown length
* 41508: gap of unknown length
* 42884: contig of 1257 bp in length
* 42885: gap of unknown length
* 42964: gap of unknown length
* 45300: contig of 2336 bp in length
* 45301: gap of unknown length
* 45400: gap of unknown length
* 46578: contig of 1178 bp in length
* 46579: gap of unknown length
* 46579: gap of unknown length
* 47749: contig of 1071 bp in length
* 47750: gap of unknown length
* 47849: gap of unknown length
* 47850: contig of 1219 bp in length
* 49069: gap of unknown length
* 49168: gap of unknown length
* 50399: contig of 1231 bp in length
* 50400: gap of unknown length
* 50499: gap of unknown length
* 51930: contig of 1431 bp in length
* 52030: gap of unknown length

```

```

* 52031: contig of 1361 bp in length
* 53391: gap of unknown length
* 53392: contig of 1439 bp in length
* 54931: gap of unknown length
* 55030: gap of unknown length
* 56380: contig of 1350 bp in length
* 56381: gap of unknown length
* 56480: contig of 1468 bp in length
* 57948: gap of unknown length
* 57949: gap of unknown length
* 58048: contig of 1036 bp in length
* 58049: gap of unknown length
* 59084: contig of 1676 bp in length
* 59185: gap of unknown length
* 60860: gap of unknown length
* 60961: contig of 1779 bp in length
* 62739: gap of unknown length
* 62840: contig of 1287 bp in length
* 64127: gap of unknown length
* 64226: contig of 1230 bp in length
* 65427: gap of unknown length
* 65456: gap of unknown length
* 65557: contig of 1045 bp in length
* 66601: gap of unknown length
* 66602: contig of 1750 bp in length
* 66702: gap of unknown length
* 68451: gap of unknown length
* 68452: gap of unknown length
* 68552: contig of 1653 bp in length
* 70205: gap of unknown length
* 70305: contig of 1024 bp in length
* 71328: gap of unknown length
* 71428: gap of unknown length
* 71429: contig of 1578 bp in length
* 73006: gap of unknown length
* 73106: gap of unknown length
* 73107: contig of 1629 bp in length
* 74335: gap of unknown length
* 74835: gap of unknown length
* 74836: contig of 1044 bp in length
* 75879: gap of unknown length
* 75880: gap of unknown length
* 75979: gap of unknown length
* 77368: gap of unknown length
* 77369: gap of unknown length
* 77468: gap of unknown length
* 78477: contig of 1009 bp in length
* 78478: gap of unknown length
* 80203: contig of 1626 bp in length
* 80204: gap of unknown length
* 80303: gap of unknown length
* 81871: contig of 1568 bp in length
* 81872: gap of unknown length
* 81971: gap of unknown length
* 83094: contig of 1123 bp in length
* 83095: gap of unknown length
* 83195: contig of 1700 bp in length
* 84894: gap of unknown length
* 84895: gap of unknown length
* 84995: contig of 1127 bp in length
* 86122: gap of unknown length
* 86221: contig of 1050 bp in length
* 87271: gap of unknown length
* 87272: contig of 1021 bp in length
* 87372: gap of unknown length
* 88392: contig of 1021 bp in length
* 88492: gap of unknown length

```

alignment\_scores:  
 Quality: 77.50 Length: 40  
 Ratio: 2.672 Gaps: 2  
 Percent Similarity: 72.500 Percent Identity: 47.500

alignment\_block:

US-09-528-682-3 x AC096092 ..

Align seg 1/1 to: AC096092 from: 1 to: 94526

```

64 MetSerThrPheGluInValProAsnAsnLysGluPheLysGlyValII 80
||||| ||| :::::::::::::: ::::: |||||
59863 ATGACACCCCTTCACAAAAGTCCCAATAACATAAATATATTGTGTGAT 59912
80 e...SerAlaAsnVallys.....TyrAspMetAsnPheLysLysLeuL 94
| :::::::::::::: ||||| ||||| ||||| |||
59913 TTTAAACACCTCAAGTGAAGATCTGTATGATGAATTTCAAGCCTTTGA 59962
94 euArgPheMetGluAspAsp 100
::: ::::::::::|||
59963 AGAAGATATTGAAGAAGAT 59982

```



seq\_name: gb\_htg:AC107529

seq\_documentation\_block:

LOCUS AC107529 135671 bp DNA linear HTG 23-JAN-2002

DEFINITION Rattus norvegicus clone CH230-15308, \*\*\* SEQUENCING IN PROGRESS

AC107529 \*\*\* 50 unordered pieces.

VERSION AC107529.1 GI:18266521

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 135671)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,  
Benton,J., Blime,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,R.,  
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,  
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
Lewis,L., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
Loisleged,H., Lozaro,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
Ogih,M., Okwuonu,G., Oragunye,N., Owiedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M.,  
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Washington,S., Williams,G., Williams,A., Wleczyk,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GKPT

Center clone name: CH230-15308

----- Summary Statistics

Sequencing vector: Plasmid; M7789

Chemistry: Dye-terminator Big Dye, 100% of reads

Assembly program: Phrap; version 0.990329First call to  
findphraplist

Consensus quality: 115465 bases at least Q40

Consensus quality: 123476 bases at least Q30

Consensus quality: 129938 bases at least Q20

Estimated insert size: 115985; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 1.6x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

consists of 50 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence.

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 7377: contig of 7377 bp in length

\* 7378 7477: gap of unknown length

\* 7478 15342: contig of 7865 bp in length

\* 15343 15442: gap of unknown length

\* 15443 22527: contig of 7085 bp in length

\* 22528 22627: gap of unknown length

\* 22628 29780: contig of 7153 bp in length

\* 29781 29880: gap of unknown length

\* 29881 35729: contig of 5849 bp in length

\* 35730 35829: gap of unknown length

\* 35830 40256: contig of 4427 bp in length

\* 40257 40356: gap of unknown length

\* 40357 43749: contig of 3393 bp in length

\* 43750 43849: gap of unknown length

\* 43850 47481: contig of 3632 bp in length

\* 47482 47581: gap of unknown length

\* 47582 51921: contig of 4340 bp in length

\* 51922 52021: gap of unknown length

\* 52022 55938: contig of 3817 bp in length

\* 55939 55938: gap of unknown length

\* 55940 55945: contig of 3907 bp in length

\* 55946 62331: gap of unknown length

\* 62332 62430: contig of 2385 bp in length

\* 62431 62439: contig of 2869 bp in length

\* 62440 65399: gap of unknown length

\* 65400 67738: contig of 2339 bp in length

\* 67739 67838: gap of unknown length

\* 67839 71179: contig of 3341 bp in length

\* 71180 71279: gap of unknown length

\* 71280 73888: contig of 2609 bp in length

\* 73889 73988: gap of unknown length

\* 73989 76228: contig of 2340 bp in length

\* 76229 76428: gap of unknown length

\* 76429 78817: contig of 2389 bp in length

\* 78818 78917: gap of unknown length

\* 78918 81814: contig of 2897 bp in length

\* 81815 83853: contig of 1939 bp in length

\* 83854 83954: gap of unknown length

\* 83955 85956: contig of 1903 bp in length

\* 85957 88030: contig of 2074 bp in length

\* 88031 88131: gap of unknown length

\* 88132 89384: contig of 1254 bp in length

\* 89385 91406: gap of unknown length

\* 91407 91506: contig of 1922 bp in length

\* 91507 94092: contig of 2586 bp in length

\* 94093 94192: gap of unknown length

\* 94193 96016: contig of 1824 bp in length

\* 96017 96117: gap of unknown length

\* 96118 98445: contig of 2229 bp in length

\* 98446 98445: gap of unknown length



```

* 98446 99777: contig of 1332 bp in length
* 99778 99877: gap of unknown length
* 101229: contig of 1352 bp in length
* 101329: gap of unknown length
* 101330 102898: contig of 1569 bp in length
* 102899 102998: gap of unknown length
* 102999 104469: contig of 1471 bp in length
* 104470 104569: gap of unknown length
* 104570 106048: contig of 1479 bp in length
* 106049 106148: gap of unknown length
* 106149 108207: contig of 2059 bp in length
* 108208 108307: gap of unknown length
* 108308 109515: contig of 1208 bp in length
* 109516 109615: gap of unknown length
* 109616 111413: contig of 1798 bp in length
* 111414 111513: gap of unknown length
* 111514 112908: contig of 1395 bp in length
* 112909 113008: gap of unknown length
* 113009 114211: contig of 1203 bp in length
* 114212 114311: gap of unknown length
* 114312 117005: contig of 2694 bp in length
* 117006 117105: gap of unknown length
* 117106 118754: contig of 1649 bp in length
* 118755 118854: gap of unknown length
* 118855 121034: contig of 2160 bp in length
* 121035 121134: gap of unknown length
* 121135 122206: contig of 1072 bp in length
* 122207 122306: gap of unknown length
* 122307 123687: contig of 1381 bp in length
* 123688 123787: gap of unknown length
* 123788 125043: contig of 1256 bp in length
* 125044 125143: gap of unknown length
* 125144 127343: contig of 2200 bp in length
* 127344 127443: gap of unknown length
* 127444 128778: contig of 1335 bp in length
* 128779 128878: gap of unknown length
* 128879 130410: contig of 1532 bp in length
* 130411 130510: gap of unknown length
* 130511 131829: contig of 1319 bp in length
* 131830 131929: gap of unknown length
* 131930 133180: contig of 1251 bp in length
* 133181 133280: gap of unknown length
* 133281 134478: contig of 1198 bp in length
* 134479 134578: gap of unknown length
* 134579 135671: contig of 1093 bp in length.

```

# FEATURES

```

source
1. .135671
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-15308"
BASE COUNT 41078 a 25388 c 25052 g 39170 t 4983 others
ORIGIN

```

```

alignment_scores:
  Quality: 77.50      Length: 40
  Ratio: 2.768       Gaps: 2
  Percent Similarity: 70.000  Percent Identity: 50.000

```

## alignment\_block:

US-09-528-682-3 x AC107529 ..

Align seg 1/1 to: AC107529 from: 1 to: 135671

```

64 MetSerThrPheGluValProAsnAsnLysGluPheLysGlyVal.. 79
||||| ||| ||||| ||||| ||||| ||||| |||||
68590 ATGACACCTTCACAAATAGTCCCAATATATAAATATCTGTGTGAC 68639

80 .ILeSerAlaAsnValLys.....TyrAspMetAsnPhelystLysLeuL 94
||||| ::||| ||||| ||||| ||||| |||||
68640 TATAAGCAAGCAAGTGAAGAGTCTGTATGACAAAGACTTCAAGTCTCTGA 68689

94 euArgPheMetGluAspAsp 100

```

```

***      :|||:|||||
68690 AGAAGAAATTGAGGAAGAT 68709
seq_name: gb_htg:AC106544

seq_documentation_block:
LOCUS      AC106544
DEFINITION Rattus norvegicus clone CH230-126B4, *** SEQUENCING IN PROGRESS
            ***, 59 unordered pieces.
ACCESSION  AC106544
VERSION    AC106544.1 GI:18139068
KEYWORDS   HTG: HTGS_PHASE1.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 163038)
AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flaeg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
            Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,O., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mahoney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
            Sisson,I., Sodergren,E., Sonaikhe,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstein,G. and Gibbs,R.
            Direct Submission
TITLE      Unpublished
JOURNAL    2 (bases 1 to 163038)
REFERENCE  Worley,K.C.
AUTHORS    Direct Submission
TITLE      Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT    ----- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ----- Project Information
            -----
            Center project name: GLID
            Center Clone name: CH230-126B4

```



```

----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 131897 bases at least Q40
Consensus quality: 140653 bases at least Q30
Consensus quality: 147742 bases at least Q20
Estimated insert size: 138591; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 59 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
*
* 1 7684: contig of 7684 bp in length
* 7685 7784: gap of unknown length
* 7785 14781: contig of 6997 bp in length
* 14782 14881: gap of unknown length
* 14882 21818: contig of 6937 bp in length
* 21819 21918: gap of unknown length
* 21919 27258: contig of 5340 bp in length
* 27259 27358: gap of unknown length
* 27359 33216: contig of 5858 bp in length
* 33217 33316: gap of unknown length
* 33317 36744: contig of 3428 bp in length
* 36745 36844: gap of unknown length
* 36845 43420: contig of 6576 bp in length
* 43421 43520: gap of unknown length
* 43521 46850: contig of 3330 bp in length
* 46851 46950: gap of unknown length
* 46951 49995: contig of 3045 bp in length
* 49996 50095: gap of unknown length
* 50096 54776: contig of 4681 bp in length
* 54777 54876: gap of unknown length
* 54877 59373: contig of 4497 bp in length
* 59374 59473: gap of unknown length
* 59474 62719: contig of 3246 bp in length
* 62720 62819: gap of unknown length
* 62820 66162: contig of 3343 bp in length
* 66163 66262: gap of unknown length
* 66263 68845: contig of 2583 bp in length
* 68846 68945: gap of unknown length
* 68946 72606: contig of 3661 bp in length
* 72607 72706: gap of unknown length
* 72707 75578: contig of 2872 bp in length
* 75579 75678: gap of unknown length
* 75679 78163: contig of 2485 bp in length
* 78164 78263: gap of unknown length
* 78264 81861: contig of 3598 bp in length
* 81862 81961: gap of unknown length
* 81962 84738: contig of 2777 bp in length
* 84739 84838: gap of unknown length
* 84839 87259: contig of 2421 bp in length
* 87260 87359: gap of unknown length
* 87360 88671: contig of 1312 bp in length
* 88672 88771: gap of unknown length
* 88772 91511: contig of 2740 bp in length
* 91512 91611: gap of unknown length
* 91612 93997: contig of 2386 bp in length
* 93998 94097: gap of unknown length
* 94098 97208: contig of 3111 bp in length
* 97209 97308: gap of unknown length
* 97309 99287: contig of 1979 bp in length
* 99288 99388: gap of unknown length
* 101242: contig of 1855 bp in length
* 101343 101343: gap of unknown length
* 101343 101343: contig of 2613 bp in length

```

```

* 103956 104055: gap of unknown length
* 104056 105716: contig of 1661 bp in length
* 105717 105816: gap of unknown length
* 105817 107247: contig of 1431 bp in length
* 107248 107347: gap of unknown length
* 107348 109665: contig of 2318 bp in length
* 109666 109765: gap of unknown length
* 109766 112602: contig of 2837 bp in length
* 112603 112702: gap of unknown length
* 112703 114809: contig of 2107 bp in length
* 114810 114909: gap of unknown length
* 114910 117201: contig of 2292 bp in length
* 117202 117301: gap of unknown length
* 117302 119479: contig of 2178 bp in length
* 119480 119579: gap of unknown length
* 119580 120781: contig of 1202 bp in length
* 120782 120881: gap of unknown length
* 120882 122921: contig of 2040 bp in length
* 122922 123021: gap of unknown length
* 123022 124039: contig of 1018 bp in length
* 124040 124139: gap of unknown length
* 124140 126304: contig of 2165 bp in length
* 126305 126404: gap of unknown length
* 126405 128986: contig of 2582 bp in length
* 128987 129086: gap of unknown length
* 129087 131429: contig of 2343 bp in length
* 131430 131529: gap of unknown length
* 131530 133834: contig of 2305 bp in length
* 133835 133934: gap of unknown length
* 133935 135695: contig of 1761 bp in length
* 135696 137408: contig of 1613 bp in length
* 137409 137508: gap of unknown length
* 137509 138895: contig of 1387 bp in length
* 138896 138995: gap of unknown length
* 138996 140666: contig of 1671 bp in length
* 140667 143419: contig of 2653 bp in length
* 143420 143519: gap of unknown length
* 143520 144655: contig of 1136 bp in length
* 144656 144755: gap of unknown length
* 144756 146793: contig of 2038 bp in length
* 146794 146893: gap of unknown length
* 146894 147964: contig of 1071 bp in length
* 147965 148064: gap of unknown length
* 148065 149748: contig of 1884 bp in length
* 149749 149848: gap of unknown length
* 149849 151701: contig of 1853 bp in length
* 151702 151801: gap of unknown length
* 151802 152870: contig of 1069 bp in length
* 152871 152970: gap of unknown length
* 152971 154909: contig of 1939 bp in length
* 154910 155009: gap of unknown length
* 155010 156334: contig of 1325 bp in length
* 156335 156434: gap of unknown length
* 156435 157578: contig of 1144 bp in length
* 157579 157678: gap of unknown length
* 157679 159130: contig of 1452 bp in length
* 159131 159230: gap of unknown length
* 159231 160825: contig of 1595 bp in length

```

```

alignment_scores:
  Quality: 77.50      Length: 40
  Ratio: 2.768      Gaps: 2
  Percent Similarity: 70.000      Percent Identity: 50.000
alignment_block:
  US-09-528-682-3 x AC106544/rev
Align seg 1/1 to reverse of: AC106544 from: 1 to: 163038
64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79
||||: ||| ||||||||| |||: |||||

```



129340 ATGACACCCCTTCAATAAGTAGTACCAAAATAATAGAAATATCTTGGTGTGAC 129291

80 .IleSerAlaAsnVallys.....TyrAspMetAsnPhelYsIleuL 94

129290 TTTAAGCAAGAGTGAAGAGTATATAGCAAGACTTCAAGTCTTTGA 129241

94 euArgPheMetGluAsp 100

129240 AGAAGAAATAGGAAGAT 129221

seq\_name: gb\_htg:AC096023

seq\_documentation\_block:

LOCUS AC096023 168120 bp DNA linear HTG 20-DEC-2001

DEFINITION Rattus norvegicus clone CH230-36K13, \*\*\* SEQUENCING IN PROGRESS

ACCESSION AC096023

VERSION AC096023.3 GI:17943680

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 168120)

REFERENCE

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbarella,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,X., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,K., Kratovic,J., Kuresha,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H., Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 168120)

Worley,K.C.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:16901704.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GEEY  
Center clone name: CH230-36K13  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329First call to findPhrapList  
Consensus quality: 131006 bases at least Q40  
Consensus quality: 140986 bases at least Q30  
Consensus quality: 150254 bases at least Q20  
Estimated insert size: 136913; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 1.8x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbankdraft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbankdraft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 69 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 5507: contig of 5507 bp in length  
5508 5607: gap of unknown length  
5608 12266: contig of 6659 bp in length  
12267 12366: gap of unknown length  
12367 18451: contig of 6085 bp in length  
18451 18451: gap of unknown length  
18452 24110: contig of 5559 bp in length  
18552 24110: gap of unknown length  
24111 27862: contig of 3652 bp in length  
27862 27862: gap of unknown length  
27863 31550: contig of 3588 bp in length  
31550 31650: gap of unknown length  
31651 34793: contig of 3143 bp in length  
34793 34793: gap of unknown length  
34794 39989: contig of 5096 bp in length  
39989 40089: gap of unknown length  
40089 43571: contig of 3482 bp in length  
43571 43571: gap of unknown length  
43572 47520: contig of 3849 bp in length  
47520 47520: gap of unknown length  
47521 50821: contig of 3201 bp in length  
50821 50921: gap of unknown length  
50921 53742: contig of 2821 bp in length  
53742 53743: gap of unknown length  
53743 57731: contig of 3889 bp in length  
57731 57732: gap of unknown length  
57732 60937: contig of 3006 bp in length  
60937 60938: gap of unknown length  
60938 63464: contig of 2527 bp in length  
63464 63465: gap of unknown length  
63465 65451: contig of 1887 bp in length  
65451 65551: gap of unknown length  
65551 65552: contig of 3944 bp in length  
65552 69595: gap of unknown length  
69595 71928: contig of 2333 bp in length  
71928 72028: gap of unknown length  
72028 72029: contig of 2578 bp in length  
72029 74607: gap of unknown length  
74607 77779: contig of 2973 bp in length  
77779 77780: gap of unknown length  
77780 80610: contig of 2831 bp in length  
80610 80710: gap of unknown length  
80710 83729: contig of 3019 bp in length  
83729 83829: gap of unknown length  
83829 86267: contig of 2438 bp in length  
86267 83830



* 86268	86367:	gap of	unknown	length
* 86368	87702:	contig of	1335	bp in length
* 87703	87802:	gap of	unknown	length
* 87803	90062:	contig of	2260	bp in length
* 90063	90162:	gap of	unknown	length
* 90163	92322:	contig of	2160	bp in length
* 92323	92422:	gap of	unknown	length
* 92423	95730:	contig of	3308	bp in length
* 95731	95830:	gap of	unknown	length
* 95831	98118:	contig of	2288	bp in length
* 98119	98218:	gap of	unknown	length
* 98219	99729:	contig of	1511	bp in length
* 99730	99829:	gap of	unknown	length
* 99830	101635:	contig of	1806	bp in length
* 101636	101735:	gap of	unknown	length
* 101736	103913:	contig of	2178	bp in length
* 103914	104013:	gap of	unknown	length
* 104014	106600:	contig of	2587	bp in length
* 106601	106700:	gap of	unknown	length
* 106701	109124:	contig of	2424	bp in length
* 109125	109224:	gap of	unknown	length
* 109225	111642:	contig of	2418	bp in length
* 111643	111742:	gap of	unknown	length
* 111743	113271:	contig of	1529	bp in length
* 113272	113371:	gap of	unknown	length
* 113372	114942:	contig of	1571	bp in length
* 114943	115042:	gap of	unknown	length
* 115043	117111:	contig of	2069	bp in length
* 117112	117211:	gap of	unknown	length
* 117212	119737:	contig of	2526	bp in length
* 119738	119837:	gap of	unknown	length
* 119838	122113:	contig of	2276	bp in length
* 122114	122213:	gap of	unknown	length
* 122214	123755:	contig of	1542	bp in length
* 123756	123855:	gap of	unknown	length
* 123856	125695:	contig of	1840	bp in length
* 125696	125795:	gap of	unknown	length
* 125796	127156:	contig of	1361	bp in length
* 127157	127256:	gap of	unknown	length
* 129257	129259:	contig of	2003	bp in length
* 129260	129359:	gap of	unknown	length
* 129360	130852:	contig of	1493	bp in length
* 130853	130952:	gap of	unknown	length
* 130953	132161:	contig of	1209	bp in length
* 132162	132261:	gap of	unknown	length
* 132262	134135:	contig of	1874	bp in length
* 134136	134235:	gap of	unknown	length
* 134236	136083:	contig of	1848	bp in length
* 136084	136183:	gap of	unknown	length
* 136184	137633:	contig of	1450	bp in length
* 137634	137733:	gap of	unknown	length
* 137734	139629:	contig of	1896	bp in length
* 139630	139729:	gap of	unknown	length
* 139730	141212:	contig of	1483	bp in length
* 141213	141312:	gap of	unknown	length
* 141313	142585:	contig of	1273	bp in length
* 142586	142685:	gap of	unknown	length
* 142686	144147:	contig of	1462	bp in length
* 144148	144247:	gap of	unknown	length
* 144248	145384:	contig of	1137	bp in length
* 145385	145484:	gap of	unknown	length
* 145485	147068:	contig of	1584	bp in length
* 147069	147168:	gap of	unknown	length
* 147169	148276:	contig of	1108	bp in length
* 148277	148376:	gap of	unknown	length
* 148377	149819:	contig of	1443	bp in length
* 149820	149919:	gap of	unknown	length

alignment\_scores: Quality: 77.50 Length: 40  
Ratio: 2.768 Gaps: 2  
Percent Similarity: 70.000 Percent Identity: 50.000

alignment\_block:  
US-09-528-682-3 x AC096023/rev ..  
  
Align seg 1/1 to reverse of: AC096023 from: 1 to: 168120  
  
64 MetSerThrPheGluGlnValProAsnAsnLysGluPhelYsGlyVal... 79  
||||: ||| ||||||||| |||: |||||  
101081 ATGACACCCCTTCATAAGTACCAATAATATAGAAATCTTGGTGTGAC 101032  
  
80 .IleSerAlaAsnVallys.....TyrAspMetAsnPhelYsLysLeuL 94  
::: ||| ::||| ||||| ||||| |||||::: |||  
101031 TTTAAGCAAGAAAGTGAAGAGTTATATGACAAAGAACTTCAAGTCTTTGA 100982  
  
94 euArgPheMetGluAspAsp 100  
::: ::|||::|||  
100981 AGAAAGAAATTGAGGAAGAT 100962  
  
seq\_name: gb\_htg:AC105576  
  
seq\_documentation\_block:  
LOCUS AC105576 214033 bp DNA linear HTG 09-JAN-2002  
DEFINITION Rattus norvegicus clone CH230-236F15, WORKING DRAFT SEQUENCE, 50  
unordered pieces.  
ACCESSION AC105576  
VERSION AC105576.1 GI:18092798  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
  
REFERENCE 1 (bases 1 to 214033)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,M., Bryant,N.P., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabist,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Hawlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseghe,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalobon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished

TITLE  
JOURNAL



REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 214033)

Worley, K.C.

Direct Submission

Submitted (09-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GNFN

Center clone name: CH230-236F15

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 185044 bases at least Q40

Consensus quality: 190296 bases at least Q30

Consensus quality: 194221 bases at least Q20

Estimated insert size: 190368; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 3.5x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

consists of 50 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

\* 1 16332: contig of 16332 bp in length

\* 16333 16432: gap of unknown length

\* 16433 26891: contig of 10459 bp in length

\* 26891 26991: gap of unknown length

\* 26992 39114: contig of 12123 bp in length

\* 39115 39214: gap of unknown length

\* 39215 47706: contig of 8492 bp in length

\* 47707 47806: gap of unknown length

\* 47807 54905: contig of 7099 bp in length

\* 54906 55005: gap of unknown length

\* 55006 61730: contig of 6725 bp in length

\* 61731 61830: gap of unknown length

\* 61831 68922: contig of 7092 bp in length

\* 68923 69022: gap of unknown length

\* 69023 73955: contig of 4933 bp in length

\* 73956 74055: gap of unknown length

\* 74056 79739: contig of 5684 bp in length

\* 79740 79839: gap of unknown length

\* 79840 87203: contig of 7364 bp in length

\* 87204 87303: gap of unknown length

\* 87304 95716: contig of 8413 bp in length

\* 95717 95816: gap of unknown length

\* 95817 101220: contig of 5404 bp in length

\* 101221 101320: gap of unknown length

\* 101321 106991: contig of 5671 bp in length

\* 106992 107091: gap of unknown length

\* 107092 111184: contig of 4093 bp in length

\* 111185 111284: gap of unknown length

\* 111285 117002: contig of 5718 bp in length

\* 117003 117103: gap of unknown length

\* 117103 122719: contig of 5617 bp in length

\* 122720 122819: gap of unknown length

\* 122820 127776: contig of 4857 bp in length

\* 127777 127776: gap of unknown length

\* 127777 132248: contig of 4472 bp in length

\* 132249 132348: gap of unknown length

\* 132349 137419: contig of 5071 bp in length

\* 137420 137519: gap of unknown length

\* 137520 140519: contig of 3000 bp in length

\* 140520 140619: gap of unknown length

\* 140620 145613: contig of 4994 bp in length

\* 145614 145713: gap of unknown length

\* 145714 150743: contig of 5030 bp in length

\* 150744 150843: gap of unknown length

\* 150844 152229: contig of 2386 bp in length

\* 152230 153329: gap of unknown length

\* 153330 155705: contig of 2376 bp in length

\* 155706 155805: gap of unknown length

\* 155806 159641: contig of 3836 bp in length

\* 159642 159741: gap of unknown length

\* 159742 162037: contig of 2196 bp in length

\* 162038 162037: gap of unknown length

\* 162038 166136: contig of 4099 bp in length

\* 166137 166236: gap of unknown length

\* 166237 169554: contig of 3418 bp in length

\* 169555 169754: gap of unknown length

\* 169755 173209: contig of 3455 bp in length

\* 173210 173309: gap of unknown length

\* 173310 174872: contig of 1563 bp in length

\* 174873 174972: gap of unknown length

\* 174973 178370: contig of 3398 bp in length

\* 178371 178470: gap of unknown length

\* 178471 181766: contig of 3296 bp in length

\* 181767 181866: gap of unknown length

\* 181867 184177: contig of 2311 bp in length

\* 184178 184277: gap of unknown length

\* 184278 186253: contig of 1976 bp in length

\* 186254 186353: gap of unknown length

\* 186354 188889: contig of 2536 bp in length

\* 188890 189899: gap of unknown length

\* 189900 190730: contig of 1741 bp in length

\* 190731 190830: gap of unknown length

\* 190831 192237: contig of 1407 bp in length

\* 192238 192337: gap of unknown length

\* 192338 194466: contig of 2129 bp in length

\* 194467 194566: gap of unknown length

\* 194567 196224: contig of 1658 bp in length

\* 196225 196324: gap of unknown length

\* 196325 197794: contig of 1470 bp in length

\* 197795 197894: gap of unknown length

\* 197895 199371: contig of 1477 bp in length

\* 199372 199471: gap of unknown length

\* 199472 202264: contig of 2793 bp in length

\* 202265 202364: gap of unknown length

\* 202365 204373: contig of 2009 bp in length

\* 204374 204473: gap of unknown length

\* 204474 205726: contig of 1253 bp in length

\* 205727 205826: gap of unknown length

\* 205827 207671: contig of 1844 bp in length

\* 207671 207770: gap of unknown length

\* 207771 208820: contig of 1050 bp in length

\* 208821 208921: gap of unknown length

\* 208922 210151: contig of 1231 bp in length

\* 210152 210251: gap of unknown length

\* 210252 211652: contig of 1401 bp in length

\* 211653 211752: gap of unknown length

\* 211753 212869: contig of 1117 bp in length

\* 212870 212969: gap of unknown length

\* 212970 214033: contig of 1064 bp in length.

FEATURES  
Location/Qualifiers  
1..214033

/organism="Rattus norvegicus"

/db\_xref="taxon:10116"

/clone="CH230-236F15"

BASE COUNT 63123 a 41277 c 42183 g 62289 t 5161 others

ORIGIN

alignment\_scores:

Quality: 77.50

Ratio: 2.981

Percent Similarity: 74.286

Length: 35

Gaps: 2

Percent Identity: 54.286



alignment\_block:

US-09-528-682-3 x AC105576 ..

Align seg 1/1 to: AC105576 from: 1 to: 214033

70 ValProAsnAsnLysGluPheLysGlyValIle...SerAlaAsnVally 85  
||||| : : : : :  
94066 GTCCCAATAACATAAAATACCTCGGTGATTTTAACCAAGCAAGTGAA 94115  
85 s.....TyrAspMetAsnPhelystLysLeuLeuArgPheMetGluAspA 100  
||||| : : : : :  
94116 AGCTGCTGATGACAGAAGACTTCAAGTCTCTGGAAGAAGAAATTTGAAGATG 94165

100 spPhe 101  
|||||

94166 ATTTC 94170

seq\_name: gb\_vi:HIVU52889

seq\_documentation\_block:

LOCUS HIVU52889 1303 bp DNA linear VRL 02-SEP-1997  
DEFINITION Human immunodeficiency virus type 1 envelope glycoprotein (env)

ACCESSION U52889

VERSION U52889.1 GI:1762238

KEYWORDS Human immunodeficiency virus type 1.

SOURCE Human immunodeficiency virus type 1

ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate

REFERENCE 1 (bases 1 to 1303)

AUTHORS Michael,N.L., Brown,A.E., Voigt,R.F., Frankel,S.S., Mascola,J.R.,

Brothers,K.S., Louder,M., Birk,D.L. and Cassol,S.A.

TITLE Rapid disease progression without seroconversion following primary

human immunodeficiency virus type 1 infection--evidence for highly

susceptible human hosts

J. Infect. Dis. 175 (6), 1352-1359 (1997)

JOURNAL 97323978

MEDLINE 2 (bases 1 to 1303)

REFERENCE Michael,N.L., Brown,A.E., Voigt,R.F., Frankel,S.S., Birk,D.L. and

Cassol,S.A.

TITLE Direct Submission

JOURNAL Submitted (28-MAR-1996) N.I. Michael, Div. of Retrovirology, Walter

Reed Army Inst. of Research, 13 Taft Court, Suite 200, Rockville,

MD 20850, USA

FEATURES Location/Qualifiers

source 1..1303

/organism="Human immunodeficiency virus type 1"

/db\_xref="taxon:11676"

/clone="patient B, clone 13"

/cell\_type="Human peripheral blood mononuclear cells"

gene 1..1303

/gene="env"

/note="V1 through fusion domain of gp160"

/codon\_start=2

/product="envelope glycoprotein"

/protein\_id="AAB68318.1"

/db\_xref="GI:1762239"

translation="WDQSLKPCVKLPLCVTLNCTDRLNATNTTSRRVEMMERGEIK

NCSNITTSIRDKVKYALPYKLDVVPIDNTINTVNNNTNINASYNRLISC

NTSVITACPKVSPFPIPIHYCAPAGFAILCKNNKNGTCPCKNVSTVQCTHIKPV

VSTQLLNSLAEEIIRSANFTDNKTIIVQLNKEQVFNKTIIVFNRRSGDPIEVM

RAFYGVTEVIGDIRAHNCISGTENWNTLSQIVEKLKEQVFNKTIIVFNRRSGDPIEVM

HTFNCGGEFFCNQTLFNSTWYKNTWDGTRVPNDITNDITLPCRKIINWQE

VKAMYPPIRGQINCSSNITGLLITRDGNNNNEIFRPGGMMRDMRSKLYKYK

VVKTEPLGVAPTRAKRVQREKRAVGIGALFLGLGAAGST"

BASE COUNT 501 a 211 c 268 g 323 t

ORIGIN

alignment\_scores:

Quality: 77.00 Length: 80

Ratio: 1.750 Gaps: 4

Percent Similarity: 55.000 Percent Identity: 28.750

alignment\_block:

US-09-528-682-3 x HIVU52889 ..

Align seg 1/1 to: HIVU52889 from: 1 to: 1303

13 ProIlleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnG1 29

||||| : : : : :  
215 CCAATAGAA...AATGATATATACTTAATGTTACTGATATAATAACTT 261

29 uValProLeuAspGlyArgTyrSerAsnPhelAlaLeuIleSerAlaGluG 46

||||| : : : : :  
262 TACTAATATTATGCCAGCTATAGCAACTAGTGTGATTAAGTTGTANCA 311

46 lyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPhe 62

||||| : : : : :  
312 CCTCAGTC.....ATTACACAAGCC...TGT 334

63 PrometSerThrPheGluGlnValPro..... 71

|||| : : : : :  
335 CCAAGGTATCTTTGAGCCAATTCACATACATATTGTGCCCGGCTGG 384

72 .....AsnAsnLysGluPheLysGly 78

385 TTTTGGCATTCCTAAGGTGTAACAATAAGAAGTTTAATGGA 424

seq\_name: gb\_vi:HIVU52894

seq\_documentation\_block:

LOCUS HIVU52894 1303 bp DNA linear VRL 02-SEP-1997  
DEFINITION Human immunodeficiency virus type 1 envelope glycoprotein (env)

ACCESSION U52894

VERSION U52894.1 GI:1762248

KEYWORDS Human immunodeficiency virus type 1.

SOURCE Human immunodeficiency virus type 1

ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate

REFERENCE 1 (bases 1 to 1303)

AUTHORS Michael,N.L., Brown,A.E., Voigt,R.F., Frankel,S.S., Mascola,J.R.,

Brothers,K.S., Louder,M., Birk,D.L. and Cassol,S.A.

TITLE Rapid disease progression without seroconversion following primary

human immunodeficiency virus type 1 infection--evidence for highly

susceptible human hosts

J. Infect. Dis. 175 (6), 1352-1359 (1997)

JOURNAL 97323978

MEDLINE 2 (bases 1 to 1303)

REFERENCE Michael,N.L., Brown,A.E., Voigt,R.F., Frankel,S.S., Birk,D.L. and

Cassol,S.A.

TITLE Direct Submission

JOURNAL Submitted (28-MAR-1996) N.I. Michael, Div. of Retrovirology, Walter

Reed Army Inst. of Research, 13 Taft Court, Suite 200, Rockville,

MD 20850, USA

FEATURES Location/Qualifiers

source 1..1303

/organism="Human immunodeficiency virus type 1"

/db\_xref="taxon:11676"

/clone="patient B, clone 4"

/cell\_type="Human peripheral blood mononuclear cells"

gene 1..1303

/gene="env"

/note="V1 through fusion domain of gp160"

/codon\_start=2

/product="envelope glycoprotein"

/protein\_id="AAB68323.1"

/db\_xref="GI:1762249"











**COMMENT**







```
80 eSerAlaasnVallysTyAspMetAsnPhelYsLysLeuLeuArgPheM 97
      :::::  |||||  |||||:::||||  :::
40802 TTTATCCAAGCAAGTGTTATGACAGAAGCTTCAAGTCTCTGAAGAAAGAA 40851

97 etGluAspAsp 100
      ::|||::|
40852 TTGAAGAAGAC 40862

seq_name: gb_pl:AF083031

seq_documentation_block:
LOCUS   AF083031                174133 bp    DNA    linear    PLN 31-OCT-2001
DEFINITION   Guillardia theta nucleomorph chromosome 3, complete sequence.
ACCESSION   AF083031
VERSION     AF083031.2  GI:13794268
KEYWORDS
SOURCE      Guillardia theta.
ORGANISM   Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
REFERENCE   1 (bases 152157 to 169088)
AUTHORS    Zauner,S., Fraunholz,M., Wastl,J., Penny,S., Beaton,M.,
Cavalier-Smith,T., Maier,U.G. and Douglas,S.
TITLE      Chloroplast protein and centrosomal genes, a tRNA intron, and odd
           telomeres in an unusually compact eukaryotic genome, the
           cryptomonad nucleomorph
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (1), 200-205 (2000)
MEDLINE    20087226
PUBMED     10618395
REFERENCE   2 (bases 1 to 174133)
AUTHORS    Douglas,S., Zauner,S., Fraunholz,M., Beaton,M., Penny,S.,
           Deng,L.T., Wu,X., Reith,M., Cavalier-Smith,T. and Maier,U.G.
TITLE      The highly reduced genome of an enslaved algal nucleus
JOURNAL    Nature 410 (6832), 1091-1096 (2001)
PUBMED     11323671
REFERENCE   3 (bases 152157 to 169088)
AUTHORS    Douglas,S.E. and Penny,S.L.
TITLE      Direct Submission
JOURNAL    Submitted (07-AUG-1998) Institute for Marine Biosciences, 1411
           Oxford Street, Halifax, NS B3H 3Z1, Canada
REFERENCE   4 (bases 1 to 174133)
AUTHORS    Zauner,S. and Douglas,S.
TITLE      Direct Submission
JOURNAL    Submitted (25-JAN-2001) Cell Biology and applied Botany,
           Philipps-University Marburg, Karl von Frisch Str, Marburg 35032,
           Germany
COMMENT    On Apr 26, 2001 this sequence version replaced gi:6690138.
FEATURES             Location/Qualifiers
     1..174133
        /organism="Guillardia theta"
        /organelle="nucleomorph"
        /db_xref="taxon:55529"
        /chromosome="3"
     279..398
        /product="5S ribosomal RNA"
        complement(936..4934)
     4934..5493
        /product="28S ribosomal RNA"
        complement(5493..5663)
     5663..6175
        /product="5.8S ribosomal RNA"
        complement(6175..8214)
     8214..9177
        /product="18S ribosomal RNA"
        /gene="orf110"
        /gene="orf110"
        /gene="orf110"
        /note="repeated at opposite end of chromosome"
        /codon_start=1
        /product="hypothetical protein"
        /protein_id="AAK39788.1"
        /db_xref="GI:13794411"
        /translation="MNKIKGKTRAGKPKKNDLTTRYTRPAPIDENRGKGTMHLSL
        PPMRSRRRAAKMIPQGGQWVLRGTSRDCKPDLPMATSSCAPISKTERLWRRWAW
        DTRFLNTS"
```

```
gene
CDS
/gene="orf210"
/701..10333
/701..10333
/gene="orf210"
/note="overlaps orf151 by 20 nucleotides at 3' end
repeated at opposite end of chromosome"
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:13794412"
/translation="NVRIFKREMGWNPQTSRGSTGHLLRILKTRHQLNNYRIQH
PIRGPSCAERIVSRNCRERTPLTKLVATRLQGITTPREWIWMSQRLSPRSYKE
EARGEDNMPRCPSPNLWTLPEEDSYNTHDSDNTRAAANLREGHQKSGSDGKPS
RPKTLSFHNQQKTRIRHLTSPQLRLEKGGTSSSRTHTLFFIVKPLGV"
complement(10314..10769)
/gene="orf151"
complement(10314..10769)
/gene="orf151"
/note="overlaps orf210 by 20 nucleotides at 3' end
repeated at opposite end of chromosome"
/codon_start=1
/product="hypothetical protein"
/translation="MRRERVCSPYFYFFLFFFFFWRESSLVFLFFFLAFHEEV
LVCFDFWRLFPPLGSGISLILNVLPLGFCGQGGSWVVERCVISFFLILSLR
GRWLDRLKCRKACPSLRGERRDSSSLSRPWVGCSTHLSISLRHPMV"
10833..11030
/gene="orf65"
10833..11030
/gene="orf65"
/note="repeated at opposite end of chromosome"
/codon_start=1
/product="hypothetical protein"
/translation="MTDRHPSCHYPSPVNFYQEKKEMTTHSSAKQVTRMPDPSKERSK
THLSQHPESLDQFTITRQGN"
11187..11396
/gene="orf69"
11187..11396
/gene="orf69"
/note="repeated at opposite end of chromosome"
/codon_start=1
/product="hypothetical protein"
/translation="MIPAAAKIPPLEKRPHPQLPARTNPGKKERWRSFTPOGRN
12264..12707
/gene="ubc4"
12264..12707
/gene="ubc4"
/note="repeated at opposite end of chromosome"
/codon_start=1
/product="ubiquitin conjugating enzyme"
/translation="WATKRIQKHDKLLKDPSPNCSAGSDENLFRWTASIIGPSDSP
YAGGVFFLSILFPDYPKPKPVQFTTKIYHPNINSNGSICLDILRDQSPALTYSKV
LLSSLLTDPNPDPDLVPEIARVFKENRSYEATAREWTQRYAT"
13171..13920
/gene="tfIID"
13171..13920
/gene="tfIID"
/note="repeated at opposite end of chromosome"
/codon_start=1
/product="TATA binding protein of transcription factor
IID"
/translation="MICSKKNSQGISQSQSIDDIILKKTNTNNTNNVFFQCKNLHAF
DVIRTRGSTEAGQSVVFRVAVRANEITPNIQNVSTVSLGIQLDLKIALKARNAEY
```







\* 1573 2850: contig of 1278 bp in length  
 \* 2851 2950: gap of unknown length  
 \* 4111: contig of 1161 bp in length  
 \* 4211: gap of unknown length  
 \* 4212: contig of 2156 bp in length  
 \* 6368 6467: gap of unknown length  
 \* 6468 7471: contig of 1004 bp in length  
 \* 7472 7571: gap of unknown length  
 \* 7572 9810: contig of 2239 bp in length  
 \* 9811 9910: gap of unknown length  
 \* 9911 14986: contig of 5076 bp in length  
 \* 14987 15086: gap of unknown length  
 \* 15087 28786: contig of 13700 bp in length  
 \* 28787 28886: gap of unknown length  
 \* 28887 39783: contig of 10897 bp in length  
 \* 39784 39883: gap of unknown length  
 \* 39884 60741: contig of 20858 bp in length  
 \* 60742 60841: gap of unknown length  
 \* 60842 78463: contig of 17622 bp in length  
 \* 78464 78563: gap of unknown length  
 \* 78564 101637: contig of 23074 bp in length  
 \* 101638 101737: gap of unknown length  
 \* 101738 114283: contig of 12546 bp in length  
 \* 114284 114383: gap of unknown length  
 \* 114384 136591: contig of 22208 bp in length  
 \* 136592 136691: gap of unknown length  
 \* 136692 193907: contig of 57216 bp in length.

# FEATURES

source  
 1..193907  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-113J6"  
 /clone\_lib="RPCI mouse BAC library 23"  
 BASE COUNT 58104 a 41331 c 40344 g 52722 t 1406 others  
 ORIGIN

## alignment\_scores:

Quality: 77.00 Length: 59  
 Ratio: 2.265 Gaps: 3  
 Percent Similarity: 57.627 Percent Identity: 35.593

## alignment\_block:

US-09-528-682-3 x AC087131 ..  
 Align seg 1/1 to: AC087131 from: 1 to: 193907  
 60 SerAsnPheProMetSerThr..... 66  
 188344 AGTCAATGGCCTTTGTCACAAAGGATAAAACAGGCTGAGAAGAAAT 188393  
 67 .....PheGluGlnValProAsnAsnLysGluPheLysG 78  
 188394 TAGGGAACACACACCTTCACAAAGGATCCCAATAAATAAAATACCTTG 188443  
 78 lyVal...IleSerAlaAsnValLys.....TyrAspMetAsnPhelys 91  
 188444 GTGTGACTCTAATCAAGAACTGAAGACCTGTATGATGAAGAACTTCAA 188493  
 92 LysLeuLeuArgPheMetGluAspAsp 100  
 188494 TCTCTGAAGAAAGAAATGAAGAAGAT 188520  
 seq\_name: gb\_ba:AP001514  
 seq\_documentation\_block:  
 LOCUS AP001514 299850 bp DNA linear BCT 10-JAN-2001  
 DEFINITION Bacillus halodurans genomic DNA, section 8/14.  
 ACCESION AP001514 BA000004  
 VERSION AP001514.1 GI:10174613  
 KEYWORDS  
 SOURCE Bacillus halodurans DNA.  
 ORGANISM Bacillus halodurans

Bacteria; Firmicutes; Bacillus/Clostridium group;  
 Bacillus/Staphylococcus group; Bacillus.  
 1 (sites)  
 Takami,H., and Horikoshi,K.  
 Identification of facultatively alkaliphilic Bacillus sp. C-125  
 to Bacillus halodurans  
 Biosci. Biotechnol. Biochem. 63, 943-945 (1999)  
 2 (sites)  
 Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G.,  
 Sasaki,R., Hiramata,C., Fuji,F. and Masui,N.  
 Genetic analysis of the chromosome of alkaliphilic Bacillus  
 halodurans C-125  
 Extremophiles 3 (3), 227-233 (1999)  
 99411980  
 3 (sites)  
 Takami,H., Nakasone,K., Hiramata,C., Takaki,Y., Masui,N., Fuji,F.,  
 Nakamura,Y. and Inoue,A.  
 An improved physical and genetic map of the genome of alkaliphilic  
 Bacillus sp. C-125  
 Extremophiles 3 (1), 21-28 (1999)  
 99184645  
 4 (sites)  
 Takami,H., Masui,N., Nakasone,K. and Horikoshi,K.  
 Replication origin region of the chromosome of alkaliphilic  
 Bacillus halodurans C-125  
 Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)  
 99356711  
 5 (sites)  
 Takami,H., Takaki,Y., Nakasone,K., Hiramata,C., Inoue,A. and  
 Horikoshi,K.  
 Sequence analysis of a 32-kb region including the major ribosomal  
 protein gene clusters from alkaliphilic Bacillus sp. strain C-125  
 Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)  
 99209008  
 6 (sites)  
 Takami,H.  
 Genome analysis of facultatively alkaliphilic Bacillus halodurans  
 C-125  
 (in) Extremophiles in deep-sea environments (Ed.):  
 HORIKOSHI, K. TSUJII:  
 : 249-284; Springer-Verlag (1999)  
 7 (sites)  
 Takami,H., Nakasone,K., Ogasawara,N., Hiramata,C., Nakamura,Y.,  
 Masui,N., Fuji,F., Takaki,Y., Inoue,A. and Horikoshi,K.  
 Sequencing of three lambda clones from the genome of alkaliphilic  
 Bacillus sp. strain C-125  
 Extremophiles 3 (1), 29-34 (1999)  
 99184646  
 8 (sites)  
 Takami,H. and Horikoshi,K.  
 Analysis of the genome of an alkaliphilic Bacillus strain from an  
 industrial point of view  
 Extremophiles 4 (2), 99-108 (2000)  
 20263314  
 9 (sites)  
 Nakasone,K., Masui,N., Takaki,Y., Sasaki,R., Maeno,G., Sakiyama,T.,  
 Hiramata,C., Fuji,F. and Takami,H.  
 Characterization and comparative study of the rrm operons of  
 alkaliphilic Bacillus halodurans C-125  
 Extremophiles 4 (4), 209-214 (2000)  
 20426005  
 10 (sites)  
 Takami,H., Nakasone,K., Takaki,Y., Maeno,G., Sasaki,R., Masui,N.,  
 Fuji,F., Hiramata,C., Nakamura,Y., Ogasawara,N., Kuhara,S. and  
 Horikoshi,K.  
 Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis  
 Nucleic Acids Res. 28 (21), 4317-4331 (2000)  
 20512582  
 11 (bases 1 to 299850)  
 Takami,H. and Takaki,Y.  
 Direct Submission  
 Submitted (22-MAR-2000) Hideto Takami, Japan Marine Science and







```
68 uGlnValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValL 85
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277105 ACAGAAGAAAACGAGAAAGG.....AAGCAAAAGTCG 277071

85 ysTyrAsp.....MetAsnPheLysLysLysLeuLeuArg 95
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277070 AGTGGGATGGAGCGCATAAAGAAAATGTCGGTAAACAACGCGTTGAG 277021

96 PheMetGluAsp 99
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277020 TGGATGGAGAT 277009

seq_name: gb_pr:AF132730

seq_documentation_block:
LOCUS AF132730 2220 bp mRNA linear PRI 01-MAY-2000
DEFINITION Homo sapiens unknown mRNA.
ACCESSION AF132730
VERSION AF132730.1 GI:7672342
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2220)
AUTHORS Luo,W.Q., Chen,J.H., Huang,X.W., Zhou,Y., Zhou,H.J., Hu,S.N. and
Yuan,J.G.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-1999) Dept. of Biochemistry, Ins. of Basic Med
Science, 5 Dong Dan 3 Tiao, Beijing 100005, People's Republic of
China
FEATURES
source Location/Qualifiers
1..2220
/organism="Homo sapiens"
/db_xref="taxon:9606"
988..2088
/note="ORI"
/codon_start=1
/product="unknown"
/db_xref="GI:7672343"
/db_xref="AAAF66441.1"
/translation="WOLHMSLTKERDQFFSEIQEIQRTSTPRDWTCKDVVAGGPER
WMLAGKNSDOLVDLLEIGSLRERKDFPGLGEGEAPAFRLFDGLVKNKPSK
KDVNLKDKMERLAEQKTFPFFNLEHRTGPSDAMAWAYTIFENIKIFHSNE
VMSQFVAVLMGRSENVYVQKETVAQLLKENTNADSONEGLLTWEQENTVLKSTFPL
KTEEQIQELMAGGHPSSNADLLNYSLFMEDEGSEPFVQKLWQYMDKDEYL
QQLKQELGELHEEVLTKRLGLMTIDPSLDKQVNTYMSOAFOLPEMPEEGDEK
EEAVVEILQALERLQVIDIRRVGREPEPAS"
BASE COUNT 626 a 550 c 640 g 404 t
ORIGIN

alignment_scores:
Quality: 76.50 Length: 129
Ratio: 1.177 Gaps: 5
Percent Similarity: 50.388 Percent Identity: 26.357

alignment_block:
US-09-528-682-3 x AF132730 ..
Align seg 1/1 to: AF132730 from: 1 to: 2220
1 AspPheThrArgAlaLeuGlnAlaTyrGluProIleGluValas 17
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1366 GATTCTCTTCAATTCTCGAGCATCGCTTGGGCCCGAGTGCAT 1415

17 nThrAsnThrValThr.....GlnIleAsnGlySerAsnG 29
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1416 GGCCTGGCGCTTACTATTTTTGAATAATCAAGATCTTCCACTCCAACG 1465

29 LuLeuAspGlyArgTyrSerAsnG 39
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
1466 AGGTTATGAGTCAGTTCTTATGCAGTCTTGTATGGGAAAGCGGAGTGAGAT 1515
40 AlaLeuIle..... 42
1516 GTGTATGTACCCAGAGGAGACAGTACGCCAGCTGCTGAAGGAGATGAC 1565
43 .....SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyT 56
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1566 AAATGCTGACAGTCAGACGAGGCGCTACTACCATGGAGCAGTTCACA 1615
56 hrValAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAsn 72
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1616 CTGCTCTCAAGATAGCTTCCCTCTCAAGACAGAGCAATCCAGGAG 1665
73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAsp....Me 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1666 CTGATGGAGCGGCGCTGGCATCCAGCAGCAGCAATGCAGACTTGCT 1715
88 tAsnPheLysLysLeuLeuArgPheMetGluAspAsp 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1716 CAACTACCGCTCACTG.....TTTATGGAGGATGAG 1746

seq_name: gb_pr:AK027245

seq_documentation_block:
LOCUS AK027245 2415 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ23592 fis, clone LNG14731, highly similar to
AF132730 Homo sapiens unknown mRNA.
ACCESSION AK027245
VERSION AK027245.1 GI:10440332
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens human lung cDNA to mRNA, clone_lib:LNG clone:LNG14731.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 2415)
AUTHORS Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo,
Laboratory of Genome Structure Analysis, Human Genome Center;
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..2415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="LNG14731"
/clone_lib="LNG"
/tissue_type="human lung"
/note="Cloning vector pME18SFL3"
misc_feature 1..2415
/note="highly similar to AF132730 Homo sapiens unknown
mRNA"
BASE COUNT 672 a 595 c 585 g 463 t
ORIGIN
```







```

* 10457 10556: gap of unknown length
* 10557 contig of 1742 bp in length
* 12398 contig of unknown length
* 12399 contig of 1295 bp in length
* 13693 contig of unknown length
* 13694 contig of 1172 bp in length
* 13794 14965: contig of unknown length
* 14966 15065: gap of unknown length
* 15066 17270: contig of 2205 bp in length
* 17271 17371: gap of unknown length
* 17371 18667: contig of 1297 bp in length
* 18668 18767: gap of unknown length
* 18768 19882: contig of 1115 bp in length
* 19883 19982: gap of unknown length
* 19983 21106: contig of 1124 bp in length
* 21107 22405: gap of unknown length
* 22406 22505: contig of 1199 bp in length
* 22506 23771: contig of 1266 bp in length
* 23772 23871: gap of unknown length
* 23872 24894: contig of 1022 bp in length
* 24894 24931: gap of unknown length
* 24931 26135: contig of 1142 bp in length
* 26136 26235: gap of unknown length
* 26236 27763: contig of 1428 bp in length
* 27764 29031: contig of 1268 bp in length
* 29032 29131: gap of unknown length
* 29132 30605: contig of 1474 bp in length
* 30606 30705: gap of unknown length
* 30706 31708: contig of 1003 bp in length
* 31709 31808: gap of unknown length
* 31809 32812: contig of 1004 bp in length.

FEATURES             Location/Qualifiers
     source            1..32812
                        /organism="Rattus norvegicus"
                        /db_xref="taxon:10116"
                        /clone="CH230-388"

BASE COUNT   9042 a   6410 c   6514 g   8436 t   2410 others
ORIGIN

alignment_scores:      Quality: 76.50      Length: 48
                       Ratio: 2.391      Gaps: 2
Percent Similarity: 66.667 Percent Identity: 41.667

alignment_block:
US-09-528-682-3 x AC094250 ..

Align seg 1/1 to: AC094250 from: 1 to: 32812
56 ThrValAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAs 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10009 ACAACCCCAAGAAAGAAATAGGGAACACACCTTCATATAATATCCCAA 10058

72 nAsnLysGluPheLysGlyVal...IleSerAlaAsnValLys.....T 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
10059 TAATATAAATACGGTGGTGACTTTATCCACGACGATTAAGATCTGT 10108

86 yrAspMetAsnPhelLysLysLeuArgPheMetGluAspAsp 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
10109 ATGATAAAGACTTCAAGCCTCTGAAGAAAGAAATGAAGAAGAC 10152

seq_name: gb_htg:AC095730

seq_documentation_block:
LOCUS      AC095730                    52562 bp    DNA        linear    HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-9A12, *** SEQUENCING IN PROGRESS ***
            32 unordered pieces.
ACCESSION  AC095730
VERSION    AC095730.2 GI:17942332
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Norway rat.

```

## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 52562)

REFERENCE  
AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J.,  
Benton, J., Blmage, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
Bowie, S., Brieva, M., Brown, E., Brown, N.P., Bryant, N.P., Buhay, C.,  
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
Duggan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,  
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,  
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,  
Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,  
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
Louiseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,  
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
Morgan, A., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,  
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,  
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N.,  
Sisson, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H.,  
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Thomas, K.,  
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S.,  
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

## TITLE

Direct Submission

2 (bases 1 to 52562)

Worley, K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15627350.

## COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GBEU

Center clone name: CH230-9A12

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 40153 bases at least Q40

Consensus quality: 46488 bases at least Q30

Consensus quality: 50909 bases at least Q20

Estimated insert size: 23609; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 0.3x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).



\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 2768: contig of 2768 bp in length  
\* 2769 2868: gap of unknown length  
\* 2869 6289: contig of 3420 bp in length  
\* 6289 6388: gap of unknown length  
\* 6389 9043: contig of 2655 bp in length  
\* 9044 9143: gap of unknown length  
\* 9144 11850: contig of 2707 bp in length  
\* 11851 11950: gap of unknown length  
\* 11951 13666: contig of 1716 bp in length  
\* 13667 13766: gap of unknown length  
\* 13767 15685: contig of 1899 bp in length  
\* 15686 15765: gap of unknown length  
\* 15766 16996: contig of 1231 bp in length  
\* 16997 17096: gap of unknown length  
\* 17097 18239: contig of 1143 bp in length  
\* 18240 18339: gap of unknown length  
\* 18340 19535: contig of 1196 bp in length  
\* 19536 19635: gap of unknown length  
\* 19636 21636: contig of 2001 bp in length  
\* 21637 21736: gap of unknown length  
\* 21737 22141: contig of 1405 bp in length  
\* 22142 2241: gap of unknown length  
\* 2242 24833: contig of 1592 bp in length  
\* 24834 24933: gap of unknown length  
\* 24934 26030: contig of 1097 bp in length  
\* 26031 26130: gap of unknown length  
\* 26131 27396: contig of 1266 bp in length  
\* 27397 27496: gap of unknown length  
\* 27497 28788: contig of 1292 bp in length  
\* 28789 28888: gap of unknown length  
\* 28889 30356: contig of 1468 bp in length  
\* 30357 30456: gap of unknown length  
\* 30457 31873: contig of 1417 bp in length  
\* 31874 31973: gap of unknown length  
\* 31974 33436: contig of 1463 bp in length  
\* 33437 33536: gap of unknown length  
\* 33537 35082: contig of 1546 bp in length  
\* 35083 35182: gap of unknown length  
\* 35183 36667: contig of 1485 bp in length  
\* 36668 36767: gap of unknown length  
\* 36768 38021: contig of 1254 bp in length  
\* 38022 38121: gap of unknown length  
\* 38122 39494: contig of 1373 bp in length  
\* 39495 39594: gap of unknown length  
\* 39595 40757: contig of 1163 bp in length  
\* 40758 40857: gap of unknown length  
\* 40858 41860: contig of 1003 bp in length  
\* 41861 41960: gap of unknown length  
\* 41961 43199: contig of 1239 bp in length  
\* 43200 43299: gap of unknown length  
\* 43300 44553: contig of 1354 bp in length  
\* 44554 44758: gap of unknown length  
\* 44759 45688: gap of unknown length  
\* 45689 45868: gap of unknown length  
\* 45869 46892: contig of 1024 bp in length  
\* 46893 46992: gap of unknown length  
\* 46993 48558: contig of 1866 bp in length  
\* 48559 48758: gap of unknown length  
\* 48759 49797: contig of 1039 bp in length  
\* 49798 49897: gap of unknown length  
\* 49898 51051: contig of 1154 bp in length  
\* 51052 51151: gap of unknown length  
\* 51152 52562: contig of 1411 bp in length.  
\* Location/Qualifiers  
\* 1. 52562

FEATURES  
source

/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-9A12"  
BASE COUNT 14699 a 11464 c 10207 g 12933 t 3259 others  
ORIGIN  
alignment\_scores:  
Quality: 76.50 Length: 41  
Ratio: 2.550 Gaps: 2  
Percent Similarity: 73.171 Percent Identity: 43.902  
alignment\_block:  
US-09-528-682-3 x AC095730 ..  
Align seg 1/1 to: AC095730 from: 1 to: 52562  
64 MetSerThrPheGluInValProAsnAsnLysGluPheLysGlyVal.. 79  
||||: |||: ||||| ||||| ||||| ||||| ||||| |||||  
5402 ATGACTCCCTTCAAAATAGTCCCAATAATATAATAATACCTCGTGTGAC 5451  
80 .lleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94  
5452 TTTAACCAAGCAAGTGAAGATCTGTATGACAAGAACTTCAAGTCATGTA 5501  
94 euArgPheMetGluAspAspPhe 101  
::: |||||: |||||  
5502 AGNATGAATTTGAGGAGATTTC 5524  
seq\_name: gb\_htg:AC095923  
seq\_documentation\_block:  
LOCUS AC095923 127014 bp DNA linear HTG 20-DEC-2001  
DEFINITION Rattus norvegicus clone CH230-10115, \*\*\* SEQUENCING IN PROGRESS  
AC095923  
AC095923.2 GI:17943560  
HTG: HTGS\_PHASE1.  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 127014)  
AUTHORS  
Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman, F.R., Allen C.,  
Alsbrooks S.L., Amaratunge H.C., Are J.R., Banks T., Barbara J.,  
Benton J., Bimaga K., Blankenburg K., Bonnin D., Bouck J.,  
Bowie S., Brieva M., Brown E., Brown M., Bryant N.P., Buha J.,  
Burch P., Burkett C., Burrell K.L., Byrd N.C., Carron T.F.,  
Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,  
Chen Z., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C.,  
Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C.,  
Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O.,  
Denn A.L., Ding Y., Dinh H.H., Douthwaite K.J., Draper H.,  
Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,  
Elhaj C., Escotto M., Falls T., Ferraguto D., Flaeg N., Ford J.,  
Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,  
Garza N., Gill R., Correll J.H., Guevara W., Gunaratne P., Hale S.,  
Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hawes A.,  
Hernandez J., Hernandez O., Hodgson A., Hoques M., Holloway C.,  
Hollins B., Homs F., Howard S., Huber J., Hulyk S., Hume J.,  
Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S.,  
Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J.,  
Kovar C., Kratovic J., Kuresh A., Landry N., Leal B., Lewis L.C.,  
Lewis L., Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W.,  
Loulseghe H., Lozano R.J., Lu X., Lucier A., Lucier R., Luna R.,  
Ma J., Maheshwari M., Mapua P., Martin R., Martindale A.,  
Martinez E., Massey E., Mawhney E., McLeod M.P., Meador M.,  
Mei G., Metzker M., Miner G., Miner Z., Mitchell T., Mohabbat K.,  
Morgan M., Morris S., Moser M., Neal D., Newton J., Newton N.,  
Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokenwo S.,  
Oguh M., Okwuonu G., Oragunye N., Oviedo R., Pace A., Payton B.,  
Peery J., Perez L., Peters L., Pickens R., Primus E., Pu L.L.,



Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,  
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

## TITLE

Direct Submission

## JOURNAL

Unpublished

2 (bases 1 to 127014)

## AUTHORS

Worley,K.C.

## TITLE

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15627543.

## COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GDWK

Center clone name: CH230-10115

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 70286 bases at least Q40

Consensus quality: 84181 bases at least Q30

Consensus quality: 96339 bases at least Q20

Estimated insert size: 76448; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 0.8x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 64 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 5855: contig of 5855 bp in length

\* 5856 5955: gap of unknown length

\* 5956 8434: contig of 2479 bp in length

\* 8435 8534: gap of unknown length

\* 8535 12540: contig of 4006 bp in length

\* 12541 12640: gap of unknown length

\* 12641 15707: contig of 3066 bp in length

\* 15707 15806: gap of unknown length

\* 15807 18783: contig of 2977 bp in length

\* 18784 18883: gap of unknown length

\* 18884 22194: contig of 3311 bp in length

\* 22195 22294: gap of unknown length

\* 22295 24691: contig of 2397 bp in length

\* 24692 24791: gap of unknown length

\* 24792 28179: contig of 3388 bp in length

\* 28180 28279: gap of unknown length

\* 28280 31510: contig of 3231 bp in length

\* 31511 31610: gap of unknown length

\* 31611 34700: contig of 3090 bp in length

\* 34701 34800: gap of unknown length

\* 34801 37701: contig of 2900 bp in length

\* 37702 37800: gap of unknown length

\* 37801 39161: contig of 1361 bp in length

\* 39162 39261: gap of unknown length

\* 39262 41956: contig of 2694 bp in length

\* 41957 42055: gap of unknown length

\* 42056 42055: gap of unknown length

42056 44199: contig of 2144 bp in length

44200 44299: gap of unknown length

44300 47149: contig of 2750 bp in length

47150 47149: gap of unknown length

47150 47150: contig of 1951 bp in length

47151 49101: gap of unknown length

49102 50694: contig of 1494 bp in length

50695 50794: gap of unknown length

50795 52845: contig of 2051 bp in length

52846 52845: gap of unknown length

52846 55010: contig of 2065 bp in length

55011 55110: gap of unknown length

55111 56646: contig of 1536 bp in length

56647 56746: gap of unknown length

56747 58981: contig of 2235 bp in length

58982 59081: gap of unknown length

59082 60996: contig of 1915 bp in length

60997 61096: gap of unknown length

61097 63506: contig of 2410 bp in length

63507 63606: gap of unknown length

63607 65170: contig of 1564 bp in length

65171 65270: gap of unknown length

65271 66508: contig of 1238 bp in length

66509 66608: gap of unknown length

66610 68633: contig of 2025 bp in length

68634 68733: gap of unknown length

68734 70884: contig of 2151 bp in length

70885 70984: gap of unknown length

70986 72197: contig of 1213 bp in length

72198 72297: gap of unknown length

72298 73971: contig of 1674 bp in length

73972 74071: gap of unknown length

74072 75960: contig of 1889 bp in length

75961 76060: gap of unknown length

76061 77462: contig of 1402 bp in length

77463 77562: gap of unknown length

77563 79550: contig of 1988 bp in length

79551 79650: gap of unknown length

79651 81406: contig of 1656 bp in length

81407 82925: contig of 1519 bp in length

82926 83025: gap of unknown length

83026 84187: contig of 1162 bp in length

84188 84287: gap of unknown length

84288 85592: contig of 1305 bp in length

85593 85692: gap of unknown length

85693 86778: contig of 1086 bp in length

86779 86878: gap of unknown length

86879 88533: contig of 1655 bp in length

88534 88633: gap of unknown length

88634 89965: contig of 1332 bp in length

89966 90065: gap of unknown length

90066 91733: contig of 1668 bp in length

91734 91833: gap of unknown length

91834 93384: contig of 1551 bp in length

93385 93485: gap of unknown length

93486 94759: contig of 1275 bp in length

94760 94859: gap of unknown length

94860 96516: contig of 1757 bp in length

96517 96716: gap of unknown length

96717 98188: contig of 1472 bp in length

98189 98288: gap of unknown length

98289 99702: contig of 1414 bp in length

99703 99802: gap of unknown length

99803 101591: contig of 1789 bp in length

101592 101691: gap of unknown length

101692 103031: contig of 1340 bp in length

103032 103131: gap of unknown length

103132 104740: contig of 1609 bp in length

104741 104840: gap of unknown length

104841 106195: contig of 1355 bp in length

106196 106295: gap of unknown length

106296 107911: contig of 1616 bp in length



```

* 107912 108011: gap of unknown length
* 108012 109017: contig of 1006 bp in length
* 109018 109117: gap of unknown length
* 109118 110541: contig of 1424 bp in length
* 110542 110641: gap of unknown length
* 110642 112107: contig of 1466 bp in length
* 112108 112207: gap of unknown length
* 112208 113580: contig of 1373 bp in length
* 113581 113680: gap of unknown length
* 113681 115276: contig of 1596 bp in length
* 115277 115376: gap of unknown length
* 115377 116635: contig of 1258 bp in length
* 116635 116734: gap of unknown length

alignment_scores:
  Quality: 76.50      Length: 42
  Ratio: 2.550        Gaps: 2
  Percent Similarity: 71.429  Percent Identity: 47.619

alignment_block:
US-09-528-682-3 x AC095923/rev ..

Align seg 1/1 to reverse of: AC095923 from: 1 to: 127014

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79
||||| |||: ||| ||||| |||: |||||
24063 ATGTCAACCTTCAAAATAGTGTGCATATATAAAATACCTTGGTGTGAC 24014

80 .IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94
||||| |||: ||| ||||| |||: |||||
24013 TCTAACCAAGCAAGTGAAGATCTGTATGACAAAGAACTTCAAGTCTCTGA 23964

94 euArgPheMetGluAspAspPheIle 102
|||: ||| |||: |||
23963 AGAAAGAAATTCAGTAGATCTCAT 23938

seq_name: gb_htg:AC106306
seq_documentation_block:
LOCUS AC106306 143318 bp DNA linear HTG 12-JAN-2002
DEFINITION Rattus norvegicus clone CH230-101122, *** SEQUENCING IN PROGRESS
***, 67 unordered pieces.
ACCESSION AC106306
VERSION AC106306.1 GI:18138827
KEYWORDS HTG; HTGS-PHASE1.
SOURCE Norway rat
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 143318)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Blum,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

```

```

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wlarczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 143318)
Worley, K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKPB
Center clone name: CH230-101122
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 114900 bases at least Q40
Consensus quality: 123396 bases at least Q30
Consensus quality: 129515 bases at least Q20
Estimated insert size: 107169; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 5282: contig of 5282 bp in length
* 5283 5382: gap of unknown length
* 5383 11647: contig of 6265 bp in length
* 11648 11747: gap of unknown length
* 11748 15927: contig of 4180 bp in length
* 15928 16027: gap of unknown length
* 16028 20427: contig of 4400 bp in length
* 20428 20527: gap of unknown length
* 20528 23839: contig of 3312 bp in length
* 23840 23939: gap of unknown length
* 23940 27026: contig of 3087 bp in length
* 27027 27126: gap of unknown length
* 27127 30418: contig of 3292 bp in length
* 30419 33966: gap of unknown length
* 33967 34067: gap of unknown length
* 34067 36590: contig of 2523 bp in length
* 36590 36690: gap of unknown length
* 36690 39337: contig of 2648 bp in length
* 39338 39338: gap of unknown length

```

## TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT



\* 39438 42213: contig of 2776 bp in length  
\* 42214 42313: gap of unknown length  
\* 42314 45728: contig of 3415 bp in length  
\* 45728 45828: gap of unknown length  
\* 45828 48978: contig of 3150 bp in length  
\* 48978 49078: gap of unknown length  
\* 49078 51402: contig of 2324 bp in length  
\* 51402 51502: gap of unknown length  
\* 51502 54348: contig of 2746 bp in length  
\* 54348 54349: gap of unknown length  
\* 54349 56770: contig of 2422 bp in length  
\* 56770 56870: gap of unknown length  
\* 56870 59114: contig of 2244 bp in length  
\* 59114 59215: gap of unknown length  
\* 59215 61337: contig of 2122 bp in length  
\* 61337 61436: gap of unknown length  
\* 61436 64073: contig of 2637 bp in length  
\* 64073 64173: gap of unknown length  
\* 64173 66450: contig of 2277 bp in length  
\* 66450 66551: gap of unknown length  
\* 66551 67794: contig of 1243 bp in length  
\* 67794 67894: gap of unknown length  
\* 67894 69558: contig of 1765 bp in length  
\* 69558 71877: contig of 2119 bp in length  
\* 71877 71977: gap of unknown length  
\* 71977 74403: contig of 2426 bp in length  
\* 74403 74503: gap of unknown length  
\* 74503 76116: contig of 1613 bp in length  
\* 76116 76216: gap of unknown length  
\* 76216 77369: contig of 1153 bp in length  
\* 77369 77469: gap of unknown length  
\* 77469 79291: contig of 1821 bp in length  
\* 79291 81563: contig of 2173 bp in length  
\* 81563 81663: gap of unknown length  
\* 81663 83185: contig of 1522 bp in length  
\* 83185 83285: gap of unknown length  
\* 83285 85338: contig of 2053 bp in length  
\* 85338 85439: gap of unknown length  
\* 85439 86906: contig of 1468 bp in length  
\* 86906 87006: gap of unknown length  
\* 87006 89351: contig of 2345 bp in length  
\* 89351 91130: contig of 1679 bp in length  
\* 91130 91230: gap of unknown length  
\* 91230 93079: contig of 1849 bp in length  
\* 93079 93179: gap of unknown length  
\* 93179 94885: contig of 1706 bp in length  
\* 94885 94985: gap of unknown length  
\* 94985 96330: contig of 1345 bp in length  
\* 96330 96431: gap of unknown length  
\* 96431 97679: contig of 1249 bp in length  
\* 97679 97779: gap of unknown length  
\* 97779 99081: contig of 1302 bp in length  
\* 99081 99181: gap of unknown length  
\* 99181 100697: contig of 1516 bp in length  
\* 100697 100797: gap of unknown length  
\* 100797 102215: contig of 1418 bp in length  
\* 102215 102315: gap of unknown length  
\* 102315 103784: contig of 1469 bp in length  
\* 103784 103884: gap of unknown length  
\* 103884 104985: contig of 1101 bp in length  
\* 104985 105085: gap of unknown length  
\* 105085 106614: contig of 1529 bp in length  
\* 106614 106714: gap of unknown length  
\* 106714 108997: contig of 2283 bp in length  
\* 108997 110453: contig of 1356 bp in length  
\* 110453 110553: gap of unknown length  
\* 110553 111733: contig of 1180 bp in length  
\* 111733 111834 113573: contig of 1740 bp in length

\* 113574 113673: gap of unknown length  
\* 113674 115059: contig of 1386 bp in length  
\* 115060 115159: gap of unknown length  
\* 115160 116359: contig of 1200 bp in length  
\* 116360 116459: gap of unknown length  
\* 116460 118315: contig of 1856 bp in length  
\* 118316 118416: gap of unknown length  
\* 118416 119813: contig of 1398 bp in length  
\* 119814 119913: gap of unknown length  
\* 119914 121334: contig of 1421 bp in length  
\* 121335 121434: gap of unknown length  
\* 121435 122983: contig of 1549 bp in length  
\* 122984 123083: gap of unknown length  
\* 123084 124931: contig of 1848 bp in length  
\* 124932 125031: gap of unknown length  
\* 125032 126077: contig of 1046 bp in length  
\* 126078 126177: gap of unknown length  
\* 126178 127400: contig of 1223 bp in length  
\* 127401 127500: gap of unknown length  
\* 127501 129156: contig of 1656 bp in length

alignment\_scores:  
Quality: 76.50 Length: 40  
Ratio: 2.732 Gaps: 2  
Percent Similarity: 70.000 Percent Identity: 50.000

alignment\_block:  
US-09-528-682-3 x AC106306/rev ...

Align seg 1/1 to reverse of: AC106306 from: 1 to: 143318

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal11 80  
||||: ||| ||||||||| :||| |||||||  
89651 ATGACACCTTTTACATAGTCCCAATAACATAATACCTTGGTGTGAT 89602

80 eSerAlaAsn...Vallys.....TyrAspMetAsnPheLysLysLeuL 94  
| :||| ||||| ||||| |||||||:|  
89601 TCTAAGCAATCAAGTGAAGAACTGTATGACAAAGAACTTCAAGTCTGTA 89552

94 euArgPheMetGluAspAsp 100  
::: |||||:|  
89551 AGAAATAATTTGAAGAGGAT 89532

seq\_name: gb\_htg:AC099071

seq\_documentation\_block:  
LOCUS AC099071 235713 bp DNA linear HTG 20-DEC-2001  
DEFINITION Rattus norvegicus chromosome Rf1 clone CH230-89J23, WORKING DRAFT  
SEQUENCE 49 unordered pieces.  
ACCESSION AC099071  
VERSION AC099071.2 GI:17943950  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 235713)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flaeg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havliak,P., Hawes,A.,



Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kravitz,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Louisgied,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
 Ogutu,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,  
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,  
 Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Watlington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs.R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GSTRZ  
 Center clone name: CH230-89J23  
 ----- Summary Statistics  
 Assembly program: Phrap: version 0.990329First call to  
 findPhrapList  
 Consensus quality: 207134 bases at least Q40  
 Consensus quality: 213086 bases at least Q30  
 Consensus quality: 217195 bases at least Q20  
 Estimated insert size: 212187; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 49 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 26072: contig of 26072 bp in length  
 \* 26073 26172: gap of unknown length  
 \* 26173 39085: contig of 12913 bp in length  
 \* 39086 39185: gap of unknown length  
 \* 39186 50587: contig of 11402 bp in length  
 \* 50588 50687: gap of unknown length  
 \* 50688 59280: contig of 8593 bp in length  
 \* 59281 59380: gap of unknown length  
 \* 59381 66584: contig of 7204 bp in length  
 \* 66585 75229: gap of unknown length  
 \* 75230 75329: gap of unknown length

75330 82704: contig of 7375 bp in length  
 82705 82804: gap of unknown length  
 82805 89750: contig of 6946 bp in length  
 89751 89850: gap of unknown length  
 89851 97348: contig of 7498 bp in length  
 97349 97448: gap of unknown length  
 104237: contig of 6789 bp in length  
 104238 104337: gap of unknown length  
 104338 110050: contig of 5713 bp in length  
 110051 110150: gap of unknown length  
 110151 115140: contig of 4990 bp in length  
 115141 115240: gap of unknown length  
 120384: contig of 5144 bp in length  
 120385 120484: gap of unknown length  
 120485 124140: contig of 3656 bp in length  
 124141 124240: gap of unknown length  
 124241 129221: contig of 5381 bp in length  
 129222 129721: gap of unknown length  
 129722 134540: contig of 4819 bp in length  
 134541 134640: gap of unknown length  
 134641 142093: contig of 7453 bp in length  
 142094 142193: gap of unknown length  
 142194 146563: gap of unknown length  
 146564 151704: contig of 5041 bp in length  
 151705 151804: gap of unknown length  
 151805 156615: contig of 4811 bp in length  
 156616 156715: gap of unknown length  
 156716 161269: contig of 4534 bp in length  
 161270 161369: gap of unknown length  
 161370 167182: contig of 5813 bp in length  
 167183 167282: gap of unknown length  
 167283 171574: contig of 4292 bp in length  
 171575 171674: gap of unknown length  
 171675 174408: contig of 2734 bp in length  
 174409 174508: gap of unknown length  
 174509 178209: contig of 3701 bp in length  
 178210 178309: gap of unknown length  
 178310 181928: contig of 3619 bp in length  
 181929 182028: gap of unknown length  
 182029 185226: contig of 3198 bp in length  
 185227 185327: gap of unknown length  
 185327 190016: contig of 4690 bp in length  
 190017 190116: gap of unknown length  
 190117 194285: contig of 4169 bp in length  
 194286 194385: gap of unknown length  
 194386 197794: contig of 3409 bp in length  
 197795 197894: gap of unknown length  
 197895 200906: contig of 3012 bp in length  
 200907 201006: gap of unknown length  
 201007 202907: contig of 1284 bp in length  
 202908 202990: gap of unknown length  
 202991 205334: contig of 2844 bp in length  
 205335 205334: gap of unknown length  
 205335 207413: contig of 2079 bp in length  
 207414 207513: gap of unknown length  
 207514 210261: contig of 2748 bp in length  
 210262 210361: gap of unknown length  
 210362 211938: contig of 1577 bp in length  
 211939 212038: gap of unknown length  
 212039 214451: contig of 2413 bp in length  
 214452 214551: gap of unknown length  
 214552 216626: contig of 2075 bp in length  
 216627 216726: gap of unknown length  
 216727 218500: contig of 1774 bp in length  
 218501 218600: gap of unknown length  
 218601 221461: contig of 2861 bp in length  
 221462 221561: gap of unknown length  
 221562 223572: contig of 2011 bp in length  
 223573 223672: gap of unknown length  
 223673 225282: contig of 1610 bp in length  
 225283 225382: gap of unknown length  
 225383 226820: contig of 1438 bp in length











Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Meador, M.,  
 Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
 Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,  
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N.,  
 Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 75754)  
 Worley, K. C.

Direct Submission  
 Submitted (23-JAN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GJAS  
 Center clone name: CH230-126B18  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; M77789  
 Chemistry: Dye-terminator Big Dye; 99% of reads  
 Assembly program: Phrap; version 0.990329 First call to  
 findPhrapList

Consensus quality: 55246 bases at least Q40  
 Consensus quality: 59652 bases at least Q30  
 Consensus quality: 64248 bases at least Q20  
 Estimated insert size: 51495; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 0.5x in Q20 bases; sum-of-contigs estimation  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 40 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 2933: contig of 2933 bp in length  
 2934 3033: gap of unknown length  
 3034 6414: contig of 3381 bp in length  
 6415 6514: gap of unknown length  
 6515 10834: contig of 4320 bp in length  
 10835 10934: gap of unknown length  
 10935 13800: contig of 2866 bp in length  
 13801 13900: gap of unknown length  
 13901 17728: contig of 3828 bp in length  
 17729 17829: gap of unknown length  
 17830 20405: contig of 2577 bp in length  
 20406 20505: gap of unknown length  
 20506 22957: contig of 2452 bp in length  
 22958 23058: gap of unknown length  
 23059 25209: contig of 2152 bp in length  
 25210 25309: gap of unknown length  
 25310 27261: contig of 1952 bp in length  
 27262 27361: gap of unknown length

27362 30057: contig of 2695 bp in length  
 30058 30156: gap of unknown length  
 30157 32059: contig of 1903 bp in length  
 32060 32159: gap of unknown length  
 32160 34232: contig of 2073 bp in length  
 34233 34332: gap of unknown length  
 34333 36000: contig of 1668 bp in length  
 36001 36100: gap of unknown length  
 36101 38066: contig of 1966 bp in length  
 38067 38166: gap of unknown length  
 38167 40154: contig of 1988 bp in length  
 40155 40254: gap of unknown length  
 40255 41630: contig of 1376 bp in length  
 41631 41730: gap of unknown length  
 41731 42881: contig of 1151 bp in length  
 42882 42981: gap of unknown length  
 42982 44123: contig of 1142 bp in length  
 44124 44223: gap of unknown length  
 44224 46149: contig of 1926 bp in length  
 46150 46249: gap of unknown length  
 46250 48397: gap of unknown length  
 48398 48496: gap of unknown length  
 48497 50042: contig of 1546 bp in length  
 50043 50142: gap of unknown length  
 50143 51727: contig of 1584 bp in length  
 51728 52921: contig of 1095 bp in length  
 52922 53021: gap of unknown length  
 53022 54044: contig of 1023 bp in length  
 54045 55548: contig of 1404 bp in length  
 55549 55648: gap of unknown length  
 55649 56821: contig of 1173 bp in length  
 56822 56921: gap of unknown length  
 56922 58069: contig of 1148 bp in length  
 58070 58169: gap of unknown length  
 58170 59394: contig of 1225 bp in length  
 59395 59494: gap of unknown length  
 59495 61112: contig of 1618 bp in length  
 61113 62213: contig of 1001 bp in length  
 62214 62313: gap of unknown length  
 62314 63872: contig of 1559 bp in length  
 63873 65176: contig of 1204 bp in length  
 65177 65276: gap of unknown length  
 65277 66330: contig of 1054 bp in length  
 66331 66430: gap of unknown length  
 66431 67891: contig of 1461 bp in length  
 67892 69444: contig of 1453 bp in length  
 69445 69544: gap of unknown length  
 69545 70757: contig of 1213 bp in length  
 70758 70857: gap of unknown length  
 70858 71922: contig of 1065 bp in length  
 71923 72022: gap of unknown length  
 72023 73192: contig of 1170 bp in length  
 73193 73292: gap of unknown length  
 73293 74603: contig of 1311 bp in length  
 74604 74703: gap of unknown length  
 74704 75754: contig of 1051 bp in length.

FEATURES  
 Location/Qualifiers  
 1..75754  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-126B18"

BASE COUNT 20531 a 14655 c 14015 g 22583 t 3970 others  
 ORIGIN

alignment\_scores:  
 Quality: 76.00 Length: 92  
 Ratio: 1.382 Gaps: 5



Percent Similarity: 59.783 Percent Identity: 31.522

## alignment\_block:

US-09-528-682-3 x AC107501 ..

Align seg 1/1 to: AC107501 from: 1 to: 75754

```

13 ProfileGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnG1 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3597 CCACGGAGGTGGACTCTAAAGT.....TCTCCCTCCCACTG 3634

29 uValProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluG 46
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3635 TGTAGAAGTCTTCAAGGTCCCGGAAATTCACATTATTTTAAGGATT 3684

46 lyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPhe 62
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
3685 CTTTTCACATATCTCGGG...TTTTTGTCTTTCCAGATGAA.....TTT 3725

63 ProMetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVa 79
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3726 GAATTTGAACCTTCACACAGTCCCAATTAATAAGTATCTTGGGGT 3775

79 l...IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysL 93
| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
3776 AACTCTAAACAACAGTGAAGATCTATATGATGAAGAAGTCTATGTC 3825

93 euLeuArgPheMetGluAspAspPhe 101
|| :||| :||| :||| :||| :||| :||| :||| :||| :|||
3826 TCAGAAAGAAAGGAAGAGACTTC 3851

```

seq\_name: gb\_htg:AC108886

## seq\_documentation\_block:

LOCUS AC108886 155423 bp DNA linear HTG 01-FEB-2002  
 DEFINITION Rattus norvegicus clone RP31-392N9 strain Brown Norway, WORKING  
 DRAFT SEQUENCE, 2 unordered pieces.

ACCESSION AC108886

VERSION AC108886.1 GI:18464055

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE Norway rat.

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 155423)

AKHER,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
 Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q.,  
 Legaspi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Maskeri,B.,  
 Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A.,  
 Stantrypop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,  
 Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
 Young,A., Zhang,L.-H. and Green,E.D.  
 NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 155423)

Green,E.D.

Direct Submission

Submitted (01-FEB-2002) NIH Intramural Sequencing Center, 8717

Grovemont Circle, Gaithersburg, MD 20877, USA

## COMMENT

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: [nisc.mouse@hori.nih.gov](mailto:nisc.mouse@hori.nih.gov)

----- Project Information

Center project name: csf

Center clone name: 392N09

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319  
 Consensus quality: 155255 bases at least Q40  
 Consensus quality: 155302 bases at least Q30  
 Consensus quality: 155311 bases at least Q20  
 Insert size: 164000; agarose-fp  
 Quality coverage: 16.23x in Q20 bases; agarose-fp  
 Quality coverage: 17.14x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 48674: contig of 48674 bp in length  
 \* 48675 48774: gap of unknown length  
 \* 48775 155423: contig of 106649 bp in length.

FEATURES  
source

1..155423  
 /organism="Rattus norvegicus"  
 /strain="Brown Norway"  
 /db\_xref="taxon:10116"  
 /clone="RP31-392N9"  
 /clone\_lib="RP31"

## misc\_feature

1..48674

/note="assembly\_fragment"

clone\_end:T7

vector\_side:right

48775..155423

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:right

BASE COUNT 41891 a 33446 c 34020 g 45965 t 101 others

## ORIGIN

## alignment\_scores:

Quality: 76.00 Length: 61

Ratio: 2.054 Gaps: 4

Percent Similarity: 60.656 Percent Identity: 37.705

## alignment\_block:

US-09-528-682-3 x AC108886 ..

Align seg 1/1 to: AC108886 from: 1 to: 155423

53 LeuPheGlyThrValAsnGlnSerAsnPhe..... 62

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 152286

TTATTTGGCTCTCTA...GAGTCTAACTCTTTAGTCTTTGTATATATT 152332

63 .....ProMetSerThrPheGluGlnValProAsnAsnLysGluP 76

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 152333

GGATAAAGCCCTATATCATGATATAGGATTGGTCCCAATATATAAAT 152382

76 heLysGlyVal...IleSerAlaAsnValLys.....TyrAspMetAsn 89

: : :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| 152383

ACCTCAGTGTACTTTAAACCAAGCAAGTTAAAGATCTGTATGATAAGAAC 152432

90 PheLysLysLeuLeuArgPheMetGluAspAsp 100

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 152433

TTCAAGCCTCTGAAGAAAGAAATGAAGAAGAC 152465

seq\_name: gb\_pr:AL136526

seq\_documentation\_block:

LOCUS AL136526

DEFINITION Human DNA sequence from clone RP11-151A6 on chromosome

13q32.21-33.3, complete sequence.

ACCESSION AL136526

VERSION AL136526.27 GI:10944529

179804 bp DNA linear PRI 02-NOV-2000



KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 179804)  
JOURNAL

COMMENT  
Dunn,M.  
Direct Submission  
Submitted (02-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Oct 21, 2000 this sequence version replaced gi:10799512.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence has been finished according to sequence map criteria  
as follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 13, constructed by the Sanger Centre Chromosome 13  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr13  
RP11-151A6 is from the library RPCT-11.1 constructed at the Roswell  
Park Cancer Institute by the group of Pieter de Jong. For further  
details see http://bacpac.med.buffalo.edu/  
VECTOR: pBAC3.6

FEATURES  
source  
Location/Qualifiers  
1..179804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="13"  
/map="g32.21-33.3"  
/clone="RP11-151A6"  
/clone-lib="RPCT-11.1"  
1..646  
/note="LIM2 repeat: matches 595. .1243 of consensus"  
648..782  
/note="LIM4 repeat: matches 4418. .4555 of consensus"  
783..1144  
/note="THS1B repeat: matches 3. .364 of consensus"  
1145..1330  
/note="LIM4 repeat: matches 4555. .4744 of consensus"  
1331..1615  
/note="AluJo repeat: matches 2. .286 of consensus"  
1616..2352  
/note="LIM4 repeat: matches 4744. .5494 of consensus"  
2353..2656  
/note="AluY repeat: matches 1. .299 of consensus"  
2657..2767  
/note="LIM4 repeat: matches 5494. .5606 of consensus"  
2887..3706  
/note="LTR28 repeat: matches 232. .1020 of consensus"  
3582..3819  
/note="LTR1 repeat: matches 1. .230 of consensus"  
4354..4541  
/note="AluJo repeat: matches 1. .187 of consensus"  
4596..4880  
/note="AluJb repeat: matches 1. .298 of consensus"  
5278..5575

repeat\_region  
/note="AluSx repeat: matches 1. .299 of consensus"  
6532..6858  
/note="LIM4 repeat: matches 5152. .5507 of consensus"  
6859..7161  
/note="AluJb repeat: matches 1. .312 of consensus"  
7162..7248  
/note="LIM4 repeat: matches 5507. .5585 of consensus"  
7629..7930  
/note="AluY repeat: matches 1. .294 of consensus"  
9437..9732  
/note="AluJb repeat: matches 1. .301 of consensus"  
9764..10067  
/note="AluSg repeat: matches 1. .303 of consensus"  
1028..10340  
/note="L2 repeat: matches 2616. .2742 of consensus"  
1055..10882  
/note="L1PA13 repeat: matches 5804. .6155 of consensus"  
10912..12259  
/note="L2 repeat: matches 1039. .2403 of consensus"  
12274..12523  
/note="AluSc repeat: matches 43. .291 of consensus"  
13269..13394  
/note="L2 repeat: matches 2555. .2677 of consensus"  
14156..14229  
/note="MIR repeat: matches 59. .129 of consensus"  
14668..14966  
/note="AluSg repeat: matches 1. .299 of consensus"  
17363..17417  
/note="MIR repeat: matches 65. .119 of consensus"  
17446..17673  
/note="MIR repeat: matches 7. .250 of consensus"  
17845..17919  
/note="AluJo repeat: matches 67. .136 of consensus"  
18498..18656  
/note="MIR repeat: matches 46. .212 of consensus"  
21868..22166  
/note="AluSg repeat: matches 10. .313 of consensus"  
22493..22791  
/note="AluSg repeat: matches 1. .299 of consensus"  
22795..22824  
/note="15 copies 2 mer tt 86% conserved"  
23490..23519  
/note="15 copies 2 mer tt 90% conserved"  
24563..24854  
/note="MSTA repeat: matches 90. .426 of consensus"  
24855..25184  
/note="MER61A repeat: matches 5. .354 of consensus"  
25189..27016  
/note="MER61-internal repeat: matches 2727. .4550 of  
consensus"  
27017..27310  
/note="AluSc repeat: matches 5. .302 of consensus"  
27311..27672  
/note="MER61-internal repeat: matches 2367. .2727 of  
consensus"  
27673..27978  
/note="AluSp repeat: matches 1. .298 of consensus"  
27979..29118  
/note="MER61-internal repeat: matches 1231. .2367 of  
consensus"  
29119..29416  
/note="AluSx repeat: matches 1. .297 of consensus"  
29417..30569  
/note="MER61-internal repeat: matches 1. .1231 of  
consensus"  
30572..30885  
/note="MER61A repeat: matches 4. .353 of consensus"  
30886..30979  
/note="MSTA repeat: matches 1. .93 of consensus"  
31021..31309  
/note="AluSx repeat: matches 1. .297 of consensus"  
31484..31790  
/note="AluSc repeat: matches 1. .307 of consensus"







*	107562	107661: gap of	100 bp
*	107662	120654: contig of	12993 bp in length
*	120655	120754: gap of	100 bp
*	120656	144638: contig of	23884 bp in length
*	120657	144638: gap of	100 bp
*	144639	144738: gap of	100 bp
*	144739	152101: contig of	7363 bp in length
*	152102	152201: gap of	100 bp
*	152202	158206: contig of	6005 bp in length
*	158207	159306: gap of	100 bp
*	159307	166363: contig of	8057 bp in length
*	166364	166463: gap of	100 bp
*	166464	175450: contig of	8987 bp in length
*	175451	175550: gap of	100 bp
*	175551	184654: contig of	9104 bp in length.

[illegible]

zhang,L. H. and Green,E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 206108)  
 Green,E.D.  
 Direct Submission  
 Submitted (19-OCT-2001) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 ----- Genome Center -----  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc\\_mouse@hgrl.nih.gov](mailto:nisc_mouse@hgrl.nih.gov)  
 ----- Project Information -----  
 Center project name: cgh  
 Center clone name: 409E06  
 ----- Summary Statistics -----  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads



Assembly program: Phrap; version 0.990319  
Consensus quality: 201836 bases at least Q40  
Consensus quality: 202988 bases at least Q30  
Consensus quality: 203594 bases at least Q20  
Insert size: 203000; agarose-fp  
Quality coverage: 11.65x in Q20 bases; agarose-fp  
Quality coverage: 11.53x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 3312: contig of 3312 bp in length  
\* 3313 3412: gap of unknown length  
\* 3413 5881: contig of 2469 bp in length  
\* 5882 5981: gap of unknown length  
\* 5982 10185: contig of 4204 bp in length  
\* 10186 10285: gap of unknown length  
\* 10286 15172: contig of 4887 bp in length  
\* 15173 15272: gap of unknown length  
\* 15273 19999: contig of 4627 bp in length  
\* 19999 19999: gap of unknown length  
\* 20000 32049: contig of 12050 bp in length  
\* 32050 32149: gap of unknown length  
\* 32150 45834: contig of 13685 bp in length  
\* 45835 45934: gap of unknown length  
\* 45935 67564: contig of 21630 bp in length  
\* 67565 67664: gap of unknown length  
\* 67665 88111: contig of 20447 bp in length  
\* 88112 88212 206108: contig of 117897 bp in length.

FEATURES		source
Location/Qualifiers		
1..206108		
/organism="Rattus norvegicus"		
/strain="Brown Norway"		
/db_xref="taxon:10116"		
/clone="RP31-409E6"		
/clone_lib="RP31"		
1..3312		
misc_feature	/note="assembly_fragment"	
misc_feature	3413..5881	
misc_feature	/note="assembly_fragment"	
misc_feature	5982..10185	
misc_feature	/note="assembly_fragment"	
misc_feature	10286..15172	
misc_feature	/note="assembly_fragment"	
misc_feature	15273..19899	
misc_feature	/note="assembly_fragment"	
misc_feature	20000..32049	
misc_feature	/note="assembly_fragment"	
misc_feature	clone_end:T7	
misc_feature	vector_side:right	
misc_feature	32150..45834	
misc_feature	/note="assembly_fragment"	
misc_feature	45935..67564	
misc_feature	/note="assembly_fragment"	
misc_feature	67665..88111	
misc_feature	/note="assembly_fragment"	
misc_feature	88212..206108	
misc_feature	/note="assembly_fragment"	
misc_feature	clone_end:SP6	
misc_feature	vector_side:left	
BASE COUNT	55913 a 45088 c 46160 g 58029 t	918 others
ORIGIN		
alignment_scores:		
Quality:	76.00	Length: 61

Ratio: 2.054 Gaps: 4  
Percent similarity: 60.656 Percent identity: 37.705  
alignment\_block:  
US-09-528-682-3 x AC097583 ..  
Align seg 1/1 to: AC097583 from: 1 to: 206108  
53 LeuphGlyThrValAsnGlnSerAsnPhe..... 62  
|||||.....: : : : :  
148725 TTATTGGCTCTCA...GAGCTCAACTTCTTTAGTCTTTGTATATATT 148771  
63 .....PrometSerThrPheGluGlnValProAsnAsnLysGluP 76  
|||||.....: : : : :  
148772 GGATAAAGCCCTATATACAGATATAGGATTGGTCCCAATAATAATAAAT 148821  
76 helysGlyVal...IleSerAlaAsnValLys.....TyrAspMetAsn 89  
: : : : : : : : : : : : : : :  
148822 ACCTCAGTGTACTTTTAAACCAAGCAAGTTAAAGATCTGTATGATAAAGAC 148871  
90 PhelLysLysLeuLeuArgPheMetGluAspAsp 100  
|||||.....: : : : :  
148872 TTCAAGCCTCTGAAGAAGAAATTTGAAGAGAC 148904  
seq\_name: gb\_htg:AC094585  
seq\_documentation\_block:  
LOCUS AC094585 210285 bp DNA linear HTG 20-DEC-2001  
DEFINITION Rattus norvegicus clone CH230-4B10, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
49 unordered pieces.  
ACCESSION AC094585  
VERSION AC094585.2 GI:17941347  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 210285)  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowie,S., Brieva,M., Brown,M., Bryant,N.P., Buhaq,C.,  
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,  
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,D.C.,  
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
Loulseghe,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,  
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokan,I., Rolfe,M.,  
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerise,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,



Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 210285)  
 Worley, K.C.  
 Direct Submission  
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:15624420.

## COMMENT

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GAYO  
 Center clone name: CH230-4B10  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 180050 bases at least Q40  
 Consensus quality: 187476 bases at least Q30  
 Consensus quality: 193628 bases at least Q20  
 Estimated insert size: 190299; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 2.7x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 49 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 8759: contig of 8759 bp in length  
 \* 8760 8859: gap of unknown length  
 \* 8860 17154: contig of 8295 bp in length  
 \* 17155 17254: gap of unknown length  
 \* 17255 26197: contig of 8943 bp in length  
 \* 26198 26297: gap of unknown length  
 \* 26298 36077: contig of 9780 bp in length  
 \* 36078 36177: gap of unknown length  
 \* 36178 45344: contig of 9167 bp in length  
 \* 45345 45444: gap of unknown length  
 \* 45445 51221: contig of 5777 bp in length  
 \* 51222 51321: gap of unknown length  
 \* 51322 59804: contig of 8483 bp in length  
 \* 59805 59904: gap of unknown length  
 \* 59905 67441: contig of 7537 bp in length  
 \* 67442 67541: gap of unknown length  
 \* 67542 75531: contig of 7990 bp in length  
 \* 75532 75631: gap of unknown length  
 \* 75632 82565: contig of 6934 bp in length  
 \* 82566 82665: gap of unknown length  
 \* 82666 89252: contig of 6587 bp in length  
 \* 89253 89352: gap of unknown length  
 \* 89353 96834: contig of 7502 bp in length  
 \* 96835 96954: gap of unknown length  
 \* 96955 103544: contig of 6590 bp in length  
 \* 103545 103644: gap of unknown length  
 \* 103645 108763: contig of 5119 bp in length  
 \* 108764 108863: gap of unknown length  
 \* 108864 113497: contig of 4634 bp in length  
 \* 113498 113597: gap of unknown length  
 \* 113598 119450: contig of 5853 bp in length

119451 119550: gap of unknown length  
 \* 119551 124302: contig of 4752 bp in length  
 \* 124303 124402: gap of unknown length  
 \* 124403 129883: contig of 5281 bp in length  
 \* 129884 129783: gap of unknown length  
 \* 129784 133492: contig of 3709 bp in length  
 \* 133493 133592: gap of unknown length  
 \* 133593 137612: contig of 4020 bp in length  
 \* 137613 137712: gap of unknown length  
 \* 137713 142261: contig of 4549 bp in length  
 \* 142262 142361: gap of unknown length  
 \* 142362 146179: contig of 3818 bp in length  
 \* 146180 146279: gap of unknown length  
 \* 146280 150559: contig of 4280 bp in length  
 \* 150560 150659: gap of unknown length  
 \* 150660 154983: contig of 4324 bp in length  
 \* 154984 155083: gap of unknown length  
 \* 155084 158865: contig of 3782 bp in length  
 \* 158866 158965: gap of unknown length  
 \* 158966 162292: contig of 3327 bp in length  
 \* 162293 162392: gap of unknown length  
 \* 162393 165046: contig of 3254 bp in length  
 \* 165047 165746: gap of unknown length  
 \* 165747 167599: contig of 1853 bp in length  
 \* 167600 167699: gap of unknown length  
 \* 167700 170062: contig of 2363 bp in length  
 \* 170063 170162: gap of unknown length  
 \* 170163 172452: contig of 2290 bp in length  
 \* 172453 172552: gap of unknown length  
 \* 172553 174441: contig of 1889 bp in length  
 \* 174442 174541: gap of unknown length  
 \* 174542 177481: contig of 2940 bp in length  
 \* 177482 177581: gap of unknown length  
 \* 177582 180454: contig of 2873 bp in length  
 \* 180455 183626: contig of 3072 bp in length  
 \* 183627 183726: gap of unknown length  
 \* 183727 186150: contig of 2424 bp in length  
 \* 186151 186250: gap of unknown length  
 \* 186251 189086: contig of 2836 bp in length  
 \* 189087 189186: gap of unknown length  
 \* 189187 191883: contig of 2697 bp in length  
 \* 191884 191983: gap of unknown length  
 \* 191984 193902: contig of 1919 bp in length  
 \* 193903 194002: gap of unknown length  
 \* 194003 195664: contig of 1662 bp in length  
 \* 195665 195764: gap of unknown length  
 \* 195765 198028: contig of 2264 bp in length  
 \* 198029 198128: gap of unknown length  
 \* 198129 199325: contig of 1197 bp in length  
 \* 199326 199425: gap of unknown length  
 \* 199426 200978: contig of 1553 bp in length  
 \* 200979 201078: gap of unknown length  
 \* 201079 202352: contig of 1174 bp in length  
 \* 202353 202352: gap of unknown length  
 \* 202353 203363: contig of 1011 bp in length  
 \* 203364 203463: gap of unknown length  
 \* 203464 205297: contig of 1834 bp in length  
 \* 205298 205397: gap of unknown length  
 \* 205398 206483: contig of 1086 bp in length  
 \* 206484 206583: gap of unknown length  
 \* 206584 207966: contig of 1383 bp in length  
 \* 207967 208066: gap of unknown length  
 \* 208067 209182: contig of 1116 bp in length  
 \* 209183 209282: gap of unknown length  
 \* 209283 210285: contig of 1003 bp in length.

FEATURES  
Source

Location/Qualifiers  
 1. .210285

/organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-4B10"

BASE COUNT 57007 a 44775 c 45986 g 57536 t 4981 others  
 ORIGIN











OM of: US-09-528-682-3 to: N\_Geneseq\_032802:\* out\_format : pfs

Date: Jun 18, 2002 7:39 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp  
-Qc/cgn2.1/USPTO.spool/US09528682/runat\_18062002\_082444\_7843/app\_query.fasta\_1.689  
-DB=N\_Geneseq\_032802 -QPMF=fastap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPOP=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=1000 -DOCALLIGN=200 -THR SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFWT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682 -CGN1\_1\_677 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-3

Query length: 107

Database: N\_Geneseq\_032802:\*

Database sequences: 1736436

Database length: 858457221

Search time (sec): 554.690000

score\_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/SID55/gcgdata/geneseq/geneseq-emb1/NA197.DAT:AAAT3069 +	16	151.26	1.24	77.00	161.26
/SID55/gcgdata/geneseq/geneseq-emb1/NA195.DAT:AAAT5623 +	32	154.10	3.10	77.00	154.10
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT3065 +	32	154.10	3.10	77.00	154.10
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT3065 +	1	164.05	0.8638	76.50	164.05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT2891 -	2	154.87	2.80	76.50	154.87
/SID55/gcgdata/geneseq/geneseq-emb1/NA191.DAT:AAQ14897 +	32	146.71	7.98	73.00	146.71
/SID55/gcgdata/geneseq/geneseq-emb1/NA196.DAT:AAAT3283 +	13	153.22	3.47	73.00	153.22
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAV49345 +	13	153.22	3.47	73.00	153.22
/SID55/gcgdata/geneseq/geneseq-emb1/NA196.DAT:AAAT3284 +	23	147.37	7.34	73.00	147.37
/SID55/gcgdata/geneseq/geneseq-emb1/NA194.DAT:AAQ32680 +	30	144.81	10.44	73.00	144.81
/SID55/gcgdata/geneseq/geneseq-emb1/NA192.DAT:AAQ31857 +	31	144.62	10.40	73.00	144.62
/SID55/gcgdata/geneseq/geneseq-emb1/NA193.DAT:AAQ42757 +	32	144.16	11.08	73.00	144.16
/SID55/gcgdata/geneseq/geneseq-emb1/NA196.DAT:AAAT31897 +	32	144.16	11.08	73.00	144.16
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAV26605 +	32	144.14	11.10	73.00	144.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAV36388 +	32	144.14	11.10	73.00	144.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAV39789 +	32	144.14	11.10	73.00	144.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAV39790 +	32	144.14	11.10	73.00	144.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA196.DAT:AAAT18200 +	33	143.86	11.51	73.00	143.86
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAG2982 +	72	150.80	4.73	72.00	150.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT64576 +	71	147.80	6.94	71.00	147.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5566 +	71	144.77	10.25	71.00	144.77
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAV00292 +	71	144.67	10.38	71.00	144.67
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT73208 +	71	144.67	10.38	71.00	144.67
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5556 +	71	144.65	10.40	71.00	144.65
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT4883 +	71	144.53	12.01	71.00	144.53
/SID55/gcgdata/geneseq/geneseq-emb1/NA194.DAT:AAQ71027 +	71	140.35	18.05	71.00	140.35
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5565 +	71	129.45	73.12	71.00	129.45
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5563 +	71	127.31	96.11	71.00	127.31
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5564 +	71	127.25	96.24	71.00	127.25
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5562 +	71	127.22	97.90	71.00	127.22
/SID55/gcgdata/geneseq/geneseq-emb1/NA193.DAT:AAQ46181 +	70	138.56	22.73	70.50	138.56
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT3001 +	70	145.82	8.96	70.00	145.82
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAV13835 +	70	140.49	17.74	70.00	140.49
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT52081 +	70	137.84	24.91	70.00	137.84
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT51819 +	70	137.60	25.69	70.00	137.60
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT4951 +	70	137.12	27.32	70.00	137.12
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT41061 +	70	136.76	28.62	70.00	136.76
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5061 +	70	136.73	28.73	70.00	136.73
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT5065 +	70	131.52	56.07	70.00	131.52

/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT1463 +	70.00	103.64	2.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT1608 -	70.00	88.15	1.5e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT1490 -	70.00	73.46	9.5e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA199683 -	70.00	61.83	4.0e+05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA199682 -	70.00	61.81	4.0e+05
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAAT27067 +	69.50	144.40	10.75
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAAT27062 +	69.50	137.74	25.23
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAAT31013 +	69.50	128.52	82.38
/SID55/gcgdata/geneseq/geneseq-emb1/NA191.DAT:AAQ14896 +	69.00	142.57	13.59
/SID55/gcgdata/geneseq/geneseq-emb1/NA196.DAT:AAQ14897 +	69.00	134.72	37.17
/SID55/gcgdata/geneseq/geneseq-emb1/NA194.DAT:AAQ45247 +	69.00	134.55	37.98
/SID55/gcgdata/geneseq/geneseq-emb1/NA194.DAT:AAQ45441 +	69.00	134.55	37.98
/SID55/gcgdata/geneseq/geneseq-emb1/NA195.DAT:AAAT5624 +	69.00	134.55	37.98
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAAT2606 +	69.00	134.28	39.32
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAAT2606 +	68.50	146.00	8.75
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT54700 +	68.50	144.87	10.11
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT54703 +	68.50	144.87	10.11
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT26702 +	68.50	139.92	19.07
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT26059 +	68.50	133.06	45.98
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT26060 +	68.50	133.06	45.98
/SID55/gcgdata/geneseq/geneseq-emb1/NA197.DAT:AAAT63072 +	68.00	140.10	18.64
/SID55/gcgdata/geneseq/geneseq-emb1/NA193.DAT:AAQ50400 +	68.00	136.76	28.61
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAT568432 +	67.50	145.32	9.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT6162 +	67.50	138.89	21.79
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT6943 +	67.50	138.03	24.31
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT26059 +	67.50	136.22	30.68
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAAT17572 +	67.50	135.04	35.69
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAAT17573 +	67.50	134.64	37.58
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT14513 +	67.50	133.26	44.83
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	67.50	133.10	45.76
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	67.50	133.04	46.09
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	67.50	133.01	46.28
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	67.50	132.89	47.00
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	67.50	131.49	56.27
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	67.00	148.65	6.23
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	66.50	132.79	47.63
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	66.50	127.84	89.90
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	66.50	127.06	99.34
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	66.50	121.12	212.71
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	66.00	143.94	11.39
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	66.00	141.47	15.64
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.50	141.50	15.57
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAAT30800 +	65.50	139.01	21.45
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	65.50	134.17	39.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	65.50	134.04	40.57
/SID55/gcgdata/geneseq/geneseq-emb1/NA197.DAT:AAAT30800 +	65.50	133.51	43.42
/SID55/gcgdata/geneseq/geneseq-emb1/NA199.DAT:AAAT30800 +	65.50	132.49	49.48
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.50	129.80	69.87
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAAT30800 +	65.50	128.07	87.24
/SID55/gcgdata/geneseq/geneseq-emb1/NA194.DAT:AAAT30800 +	65.50	127.19	97.63
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.50	123.87	149.55
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.50	122.77	172.16
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAAT30800 +	65.50	122.66	174.67
/SID55/gcgdata/geneseq/geneseq-emb1/NA198.DAT:AAAT30800 +	65.00	128.17	86.09
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAT30800 +	65.00	110.12	87.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.00	108.40	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.00	108.40	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	65.00	108.40	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	64.50	137.85	24.87
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	64.50	137.60	25.71
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	64.50	136.28	30.42
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT30800 +	64.50	123.97	147.63
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	64.50	121.57	200.68
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT30800 +	64.50	121.36	206.31







/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:AA81505 +	60.00	116.98	361.77	59.50	108.85	1.0e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2989 +	60.00	116.72	373.86	59.50	108.85	1.0e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF30894 +	60.00	112.72	624.95	59.50	108.38	1.1e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF93778 +	60.00	112.15	672.18	59.50	108.32	1.1e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA70259 +	60.00	108.37	1.1e+03	59.50	108.94	1.3e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2988 -	60.00	107.69	1.2e+03	59.50	106.60	1.4e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA78868 +	60.00	106.96	1.3e+03	59.50	105.81	1.5e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA75235 +	60.00	106.56	1.4e+03	59.50	103.95	1.9e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV74277 +	60.00	106.54	1.4e+03	59.50	102.26	2.4e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA78867 +	60.00	105.53	1.6e+03	59.50	97.93	4.2e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA79882 -	60.00	101.73	2.6e+03	59.50	92.88	8.0e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ81783 +	59.50	133.26	44.81	59.50	84.22	2.4e+04
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA71486 -	59.50	131.86	53.68	59.50	70.21	1.4e+05
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK45829 -	59.50	131.86	53.68	59.00	130.67	62.48
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI1754 -	59.50	131.86	53.68	59.00	125.18	126.43
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK45049 -	59.50	131.07	59.36	59.00	124.00	147.06
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA79882 -	59.50	129.86	71.12	59.00	123.60	154.75
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH28923 +	59.50	128.93	78.11	59.00	119.20	171.96
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	128.86	78.83	59.00	118.20	309.32
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	117.69	330.20
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	117.33	345.78
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	117.04	358.85
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	116.67	376.27
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	116.66	376.27
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	127.31	96.15	59.00	116.24	397.86
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA793671 +	59.50	124.98	129.67	59.00	114.70	484.88
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2971 +	59.50	124.67	134.83	59.00	113.81	542.99
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	123.73	152.12	59.00	113.18	589.14
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	123.52	156.41	59.00	112.91	609.84
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	123.49	157.06	59.00	112.74	622.82
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	121.26	208.98	59.00	111.17	762.04
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.29	236.51	59.00	111.01	778.29
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.18	240.10	59.00	110.55	825.36
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	120.18	240.10	59.00	110.17	866.41
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	120.18	240.10	59.00	110.13	870.50
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.18	240.10	59.00	110.02	883.19
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.18	240.10	59.00	109.74	915.86
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.18	240.10	59.00	109.59	932.87
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	120.18	240.10	59.00	109.19	982.08
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AA793671 +	59.50	118.83	285.45	59.00	109.06	999.39
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA793671 +	59.50	118.83	285.45	59.00	109.05	1.0e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA793671 +	59.50	118.83	285.45	59.00	108.61	1.1e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA793671 +	59.50	118.83	285.45	59.00	107.58	1.2e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	118.83	285.45	59.00	107.35	1.2e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	118.83	285.45	59.00	107.17	1.3e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	118.83	285.45	59.00	107.12	1.3e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	118.83	285.45	59.00	107.02	1.3e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	118.83	285.45	59.00	106.38	1.4e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	117.60	333.99	59.00	106.29	1.4e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	117.48	339.47	59.00	105.46	1.6e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	103.74	2.0e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	101.24	2.7e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	98.04	4.1e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	97.92	4.2e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	97.58	4.4e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	93.75	7.1e+03
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA793671 +	59.50	116.17	401.44	59.00	86.51	1.8e+04
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA793671 +	59.50	114.76	480.69	59.00	81.36	3.5e+04
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA793671 +	59.50	114.76	480.69	59.00	77.30	5.8e+04
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA793671 +	59.50	112.39	504.21	59.00	61.05	4.4e+05
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2989 +	59.50	112.67	628.64	58.50	132.40	50.06
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2989 +	59.50	111.95	689.12	58.50	131.93	53.17
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2989 +	59.50	111.92	692.55	58.50	130.37	65.81
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AA793671 +	59.50	111.46	734.42	58.50	130.16	66.69
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	110.40	841.53	58.50	129.49	72.72
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA793671 +	59.50	108.94	1.0e+03	58.50	128.70	80.48







/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF91874 + 57.50 109.28 971.32 + 57.00 114.59 491.23 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF27437 + 57.50 109.09 995.42 + 57.00 113.72 549.62 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU27240 + 57.50 108.96 1.0e+03 + 57.00 113.72 549.62 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU13178 + 57.50 107.43 1.2e+03 + 57.00 113.52 563.53 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF29935 + 57.50 107.40 1.2e+03 + 57.00 113.02 601.40 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF63954 + 57.50 106.68 1.4e+03 + 57.00 113.02 601.40 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH13926 + 57.50 106.52 1.4e+03 + 57.00 112.74 622.77 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH43350 + 57.50 106.52 1.4e+03 + 57.00 112.74 622.77 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH569876 + 57.50 106.47 1.4e+03 + 57.00 111.99 686.27 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAU02948 + 57.50 106.44 1.4e+03 + 57.00 111.99 686.27 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH45411 + 57.50 106.44 1.4e+03 + 57.00 111.99 686.27 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH389125 + 57.50 106.18 1.4e+03 + 57.00 110.78 801.27 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH85341 + 57.50 106.16 1.4e+03 + 57.00 110.76 803.62 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH63990 + 57.50 105.82 1.5e+03 + 57.00 110.76 803.62 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH587118 + 57.50 105.71 1.5e+03 + 57.00 110.28 854.44 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAU03967 + 57.50 105.69 1.5e+03 + 57.00 110.24 858.61 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAU122820 + 57.50 105.33 1.6e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU11009 + 57.50 105.20 1.6e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU112846 + 57.50 105.17 1.6e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH32926 + 57.50 104.98 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH45582 + 57.50 104.77 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH54402 + 57.50 104.74 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAU75487 + 57.50 104.69 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAU32748 + 57.50 104.69 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAU1797 + 57.50 104.69 1.7e+03 + 57.00 110.12 872.35 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAU78246 + 57.50 104.46 1.8e+03 + 57.00 109.88 898.79 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAU13732 + 57.50 104.39 1.8e+03 + 57.00 109.77 911.48 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABK09693 + 57.50 104.29 1.8e+03 + 57.00 109.15 987.90 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU12128 + 57.50 103.76 2.0e+03 + 57.00 108.68 1.0e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH54209 + 57.50 103.74 2.0e+03 + 57.00 108.57 1.1e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU27917 + 57.50 103.65 2.0e+03 + 57.00 108.56 1.1e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH54816 + 57.50 103.58 2.0e+03 + 57.00 107.83 1.2e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAU78245 + 57.50 101.55 2.6e+03 + 57.00 107.83 1.2e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU12294 + 57.50 100.99 2.8e+03 + 57.00 106.91 1.3e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU18232 + 57.50 100.97 2.8e+03 + 57.00 106.91 1.3e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABU11008 + 57.50 99.94 3.0e+03 + 57.00 106.59 1.4e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU11008 + 57.50 99.94 3.0e+03 + 57.00 106.59 1.4e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH63315 + 57.50 99.47 3.4e+03 + 57.00 106.15 1.5e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU12576 + 57.50 99.47 3.4e+03 + 57.00 106.15 1.5e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU13616 + 57.50 97.81 4.2e+03 + 57.00 105.80 1.5e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAU52282 + 57.50 91.94 9.0e+03 + 57.00 105.27 1.6e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU1396 + 57.50 90.51 1.1e+04 + 57.00 105.14 1.7e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU03551 + 57.50 90.42 1.1e+04 + 57.00 104.82 1.7e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU27214 + 57.50 89.81 1.2e+04 + 57.00 104.77 1.7e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU13680 + 57.50 89.46 1.2e+04 + 57.00 104.09 1.9e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU14580 + 57.50 89.46 1.2e+04 + 57.00 104.08 1.9e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH81486 + 57.50 83.35 2.7e+04 + 57.00 104.08 1.9e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU06898 + 57.50 81.65 3.4e+04 + 57.00 104.08 1.9e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAU66772 + 57.50 80.63 3.8e+04 + 57.00 103.97 1.9e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAU1544 + 57.50 79.28 4.5e+04 + 57.00 103.63 2.0e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAU86431 + 57.50 57.36 6.8e+05 + 57.00 103.33 2.1e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68529 + 57.50 57.36 6.8e+05 + 57.00 101.96 2.5e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH81490 + 57.50 42.67 2.6e+06 + 57.00 101.89 2.5e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH56682 + 57.00 123.78 151.25 + 57.00 101.67 2.6e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAU6613 + 57.00 122.88 169.67 + 57.00 101.61 2.6e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAU75116 + 57.00 121.03 215.18 + 57.00 101.31 2.7e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAU71852 + 57.00 120.68 224.93 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH96114 + 57.00 120.68 224.93 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAU71851 + 57.00 118.58 294.74 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH596113 + 57.00 118.58 294.74 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU11801 + 57.00 118.55 295.65 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAU3623 + 57.00 118.47 298.86 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAU70734 + 57.00 118.45 299.77 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAU42220 + 57.00 118.45 299.77 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAU79599 + 57.00 118.45 299.77 + 57.00 101.03 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU11893 + 57.00 116.48 385.48 + 57.00 100.92 2.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAU12765 + 57.00 116.41 389.39 + 57.00 99.86 3.3e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAU3622 + 57.00 116.21 399.20 + 57.00 99.83 3.3e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68456 + 57.00 115.50 437.56 + 57.00 98.68 3.8e+03 +  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU18899 + 57.00 115.20 454.26 + 57.00 97.88 4.2e+03 +



/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL22106 - 57.00 97.58 4.4e+03 6  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:ABL13147 - 57.00 96.38 5.1e+03 72  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL22628 + 57.00 94.86 6.2e+03 8  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL08598 + 57.00 94.73 6.3e+03 8  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABL32462 - 57.00 94.00 6.9e+03 91  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL03530 + 57.00 90.47 1.1e+04 1  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV62176 - 57.00 67.50 2.0e+05 11  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAV25119 - 57.00 64.61 2.9e+05 15  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH96431 - 57.00 56.13 7.8e+05 3  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68532 - 57.00 56.13 7.8e+05 3  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV58840 + 57.00 50.98 1.4e+06 58  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV20248 + 57.00 46.19 2.1e+06 9  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV22305 - 57.00 44.39 2.4e+06 10  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI99683 + 57.00 29.80 3.5e+06 4  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI99682 + 57.00 29.79 3.5e+06 4  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC0508 - 56.50 125.24 125.45 40  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC34375 - 56.50 132.39 183.03 55  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ51575 - 56.50 120.91 218.50 61  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV73051 + 56.50 120.74 223.28 63  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ51576 + 56.50 120.58 228.08 63  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ54551 - 56.50 120.56 228.56 63  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAZ52266 + 56.50 120.25 237.77 66  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ54271 - 56.50 119.47 262.92 70  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV51218 + 56.50 117.45 340.37 8  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV53229 + 56.50 116.38 390.79 9  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24769 + 56.50 116.12 404.08 96  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV75784 + 56.50 116.12 404.08 96  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ65933 + 56.50 115.43 441.18 10  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH81316 + 56.50 115.43 441.18 10  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH13346 + 56.50 114.88 473.29 10  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABAO05817 + 56.50 114.01 529.61 11  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ29011 + 56.50 113.95 533.78 11  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ54516 + 56.50 112.97 604.62 13  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV24826 + 56.50 112.97 604.62 13  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ38917 - 56.50 112.09 676.98 14  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ97486 + 56.50 112.09 676.98 14  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ23551 + 56.50 111.93 690.93 14  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ63375 + 56.50 111.93 690.93 14  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ42745 - 56.50 111.62 719.69 14  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV47072 + 56.50 111.31 748.73 15  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAV70539 + 56.50 111.09 779.37 15  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV32228 + 56.50 110.95 783.96 15  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74707 + 56.50 110.38 843.56 16  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV26704 + 56.50 110.38 843.56 16  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH67644 + 56.50 110.12 859.63 16  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV29669 + 56.50 110.12 871.74 17  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV09394 - 56.50 109.96 889.98 17  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV16635 + 56.50 109.82 894.72 17  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL11637 + 56.50 109.81 906.95 17  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV26704 + 56.50 109.71 918.54 18  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAV1110 + 56.50 109.66 924.68 18  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV67752 + 56.50 109.54 939.07 18  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV67752 + 56.50 109.51 943.19 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV55358 + 56.50 109.21 980.48 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74272 + 56.50 109.16 986.73 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV7063 + 56.50 109.16 986.73 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ88066 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ95031 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV02471 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ21078 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ59635 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV7595 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV44843 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV7849 + 56.50 109.01 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV29678 + 56.50 108.80 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV60621 + 56.50 108.77 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAV07341 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV09764 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV26650 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV21818 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV08975 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV67037 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV12047 + 56.50 108.75 1.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19647 + 56.50 108.40 1.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAV70692 + 56.50 108.10 1.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV72029 + 56.50 108.10 1.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV50959 + 56.50 107.77 1.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV24954 - 56.50 107.77 1.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV74269 - 56.50 107.54 1.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74269 - 56.50 107.35 1.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV77052 + 56.50 107.35 1.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV29672 + 56.50 106.93 1.3e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL14527 + 56.50 106.46 1.4e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV19464 - 56.50 106.42 1.4e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV19464 - 56.50 105.83 1.5e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV91572 + 56.50 105.83 1.5e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL22650 + 56.50 105.73 1.5e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV55168 + 56.50 105.12 1.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL25602 - 56.50 104.97 1.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV79266 + 56.50 104.84 1.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19889 + 56.50 103.50 2.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL03599 + 56.50 103.37 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV56991 + 56.50 103.32 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV34620 - 56.50 103.26 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV15174 - 56.50 103.26 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH76457 - 56.50 103.26 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV71209 + 56.50 103.25 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV88712 + 56.50 103.13 2.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV16517 + 56.50 102.91 2.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV38478 + 56.50 102.91 2.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV58587 + 56.50 102.85 2.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAV70691 + 56.50 102.82 2.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL04218 + 56.50 102.81 2.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV27028 + 56.50 102.60 2.3e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL23458 - 56.50 102.35 2.4e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV38466 + 56.50 101.73 2.6e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV23700 + 56.50 101.72 2.6e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV82995 + 56.50 101.72 2.6e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV68658 + 56.50 101.20 2.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAV7791 + 56.50 101.03 2.8e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV85767 + 56.50 101.03 2.8e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAV71111 + 56.50 100.90 2.8e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV60173 - 56.50 100.57 3.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV60174 - 56.50 100.57 3.0e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19646 + 56.50 99.78 3.3e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52212 - 56.50 98.92 3.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL13935 + 56.50 98.78 3.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL14526 + 56.50 98.55 3.8e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19138 - 56.50 97.89 4.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV29882 + 56.50 97.86 4.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV29882 + 56.50 97.80 4.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ10370 + 56.50 97.74 4.3e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19888 - 56.50 97.56 4.4e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV13218 + 56.50 97.46 4.4e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL03598 - 56.50 97.09 4.6e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV29673 + 56.50 96.93 4.7e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL19136 - 56.50 96.74 4.8e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV40090 + 56.50 96.39 5.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV49096 + 56.50 96.37 5.1e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABL32498 - 56.50 94.76 6.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV87953 + 56.50 93.67 7.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV13934 + 56.50 93.67 7.2e+03 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV13934 + 56.50 88.50 1.4e+04 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV09525 + 56.50 88.30 1.4e+04 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV73236 + 56.50 87.92 1.5e+04 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL17076 - 56.50 80.57 3.6e+04 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV34005 + 56.50 73.51 9.4e+04 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV281210 - 56.50 72.08 1.1e+05 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV28542 - 56.50 43.59 2.5e+06 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV22305 + 56.50 43.16 2.6e+06 19  
/SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV01545 + 56.50 121.76 195.86 19



/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAT33925 -	56.00	116.83	368.96	80	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA52342 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ53934 +	56.00	116.01	409.88	87	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA22137 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ54471 +	56.00	116.01	409.88	87	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK00614 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS51685 +	56.00	115.72	425.03	8	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK26084 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV54789 +	56.00	115.62	430.74	9	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK10692 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74842 +	56.00	115.18	455.72	94	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK131948 -	55.50	121.60	199.91
/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ22702 +	56.00	114.59	491.54	99	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK68317 +	55.50	120.02	244.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAE91418 +	56.00	114.56	493.51	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK68814 +	55.50	120.26	269.97
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC55025 +	56.00	113.60	558.34	10	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK61550 +	55.50	118.53	296.40
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC39133 +	56.00	113.58	559.70	10	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK61549 +	55.50	117.65	332.13
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAC0406 -	56.00	113.47	567.18	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK34077 +	55.50	117.46	340.20
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABA90337 -	56.00	113.47	567.18	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK12404 +	55.50	117.05	358.55
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH04004 -	56.00	113.31	579.47	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK39897 +	55.50	116.46	386.86
/SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAH70101 +	56.00	110.67	812.14	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAK7429 +	55.50	115.75	423.52
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK61600 +	56.00	110.53	827.08	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK33426 +	55.50	113.65	554.19
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ56308 +	56.00	110.49	831.57	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK53931 +	55.50	113.37	574.88
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK53931 +	56.00	110.34	847.35	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK52683 +	55.50	112.37	653.09
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK61699 +	56.00	110.16	867.75	15	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK527504 +	55.50	112.27	661.84
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH01053 +	56.00	108.24	976.18	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH53998 +	55.50	111.73	709.28
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH33297 +	56.00	108.71	1.0e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK51961 +	55.50	111.28	751.76
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU20527 +	56.00	108.53	1.1e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK13492 +	55.50	111.12	767.44
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK51718 -	56.00	108.02	1.1e+03	16	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK13441 -	55.50	111.07	772.41
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK583069 +	56.00	107.48	1.2e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA69639 -	55.50	110.95	784.03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ14479 +	56.00	106.91	1.3e+03	20	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK54552 +	55.50	110.85	794.02
/SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:AAK08922 -	56.00	106.59	1.4e+03	21	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK33558 +	55.50	110.65	814.94
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK572365 +	56.00	104.30	1.8e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK53672 -	55.50	110.37	844.46
/SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAK17303 +	56.00	104.19	1.9e+03	27	/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ81780 +	55.50	110.12	871.68
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV20201 +	56.00	104.16	1.9e+03	27	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK52357 +	55.50	110.12	871.68
/SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ14475 +	56.00	103.40	2.1e+03	29	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24781 +	55.50	110.02	882.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAK50644 +	56.00	103.36	2.1e+03	29	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAK68217 +	55.50	110.02	882.80
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH02915 +	56.00	103.30	2.2e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAK14426 -	55.50	109.87	930.78
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK592967 +	56.00	103.01	2.2e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA82980 +	55.50	109.54	939.78
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK74625 +	56.00	102.74	2.2e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK37852 +	55.50	109.51	942.39
/SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ73755 +	56.00	102.45	2.3e+03	32	/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ34689 +	55.50	109.47	947.62
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK572321 +	56.00	102.04	2.5e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK13456 +	55.50	109.15	987.06
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK57062 +	56.00	101.67	2.6e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ85986 +	55.50	108.72	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:AAK70681 +	56.00	101.65	2.6e+03	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK42127 +	55.50	108.41	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ14478 +	56.00	101.64	2.6e+03	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK29149 +	55.50	108.25	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ48497 +	56.00	101.62	2.6e+03	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH4575 +	55.50	107.79	1.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAK93651 +	56.00	101.62	2.6e+03	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV33033 +	55.50	106.79	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:AAK50364 +	56.00	101.12	2.8e+03	36	/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV33017 +	55.50	106.79	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK10475 +	56.00	101.12	2.8e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV27433 +	55.50	106.77	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK585083 +	56.00	101.12	2.8e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK66071 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH81312 +	56.00	100.20	3.1e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK66072 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:AAK50516 +	56.00	100.16	3.1e+03	39	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK77803 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV45523 +	56.00	99.59	3.4e+03	42	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK77804 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV59093 +	56.00	98.15	4.0e+03	48	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK61190 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK85760 -	56.00	97.90	4.2e+03	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK61191 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK80231 +	56.00	97.70	4.3e+03	50	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK61192 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK25648 +	56.00	97.44	4.4e+03	5	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK61193 +	55.50	106.76	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK57181 +	56.00	96.59	4.9e+03	56	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU20901 +	55.50	105.33	1.6e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU33580 -	56.00	92.08	8.8e+03	86	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU26637 +	55.50	105.29	1.6e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABU32808 -	56.00	91.51	9.5e+03	9	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK61058 +	55.50	105.25	1.6e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK90888 -	56.00	89.62	1.2e+04	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24873 +	55.50	104.40	1.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAK22335 +	56.00	85.18	2.1e+04	16	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK64209 +	55.50	103.98	1.9e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK1515 -	56.00	83.81	2.5e+04	19	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK67957 +	55.50	102.63	2.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52168 -	56.00	81.26	3.5e+04	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ50430 +	55.50	101.28	2.7e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK78367 +	56.00	81.26	3.5e+04	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU12861 +	55.50	101.20	2.7e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK84723 +	56.00	80.49	3.9e+04	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK79271 +	55.50	101.11	2.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK19696 -	56.00	75.78	7.1e+04	41	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH54916 -	55.50	100.18	3.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK68247 -	56.00	75.78	7.1e+04	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK76918 +	55.50	99.60	3.4e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK682106 -	56.00	75.44	7.4e+04	43	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU18275 +	55.50	99.16	3.6e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK68254 -	56.00	68.98	1.7e+05	8	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU20985 +	55.50	98.83	3.7e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAK18650 -	56.00	66.73	2.2e+05	9	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU19577 +	55.50	98.56	3.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK28500 +	56.00	53.67	1.0e+06	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU19575 +	55.50	98.14	4.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK21609 +	56.00	36.47	1.3e+06	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ25536 +	55.50	98.01	4.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAK24063 +	56.00	36.47	1.3e+06	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ25534 +	55.50	98.00	4.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABA03041 -	56.00	31.53	3.5e+06	29	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ52047 +	55.50	98.00	4.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK1474 +	55.50	127.17	97.90	26	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAZ52089 +	55.50	97.87	4.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK102311 +	55.50	125.64	119.19	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ52048 +	55.50	97.80	4.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24655 +	55.50	125.43	122.32	31	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV25095 +	55.50	97.80	4.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ81782 -	55.50	122.34	181.80	42	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU11793 +	55.50	96.79	4.8e+03



```

/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAQ64556 - 55.50 95.28 5.9e+03 58
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AA13390 + 55.50 95.00 6.1e+03 58
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV52200 - 55.50 92.86 8.0e+03 71
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL18274 + 55.50 92.58 8.3e+03 77
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL20984 + 55.50 92.40 8.5e+03 77
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL19576 + 55.50 92.10 8.8e+03 77
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ36030 - 55.50 91.82 9.1e+03 79
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA126526 - 55.50 91.35 9.7e+03 8
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA126553 - 55.50 91.35 9.7e+03 8
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA126600 - 55.50 91.35 9.7e+03 8
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ42541 + 55.50 90.84 1.0e+04 87

```

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAF63069

seq\_documentation\_block:

ID AAT63069 standard; cDNA; 1622 BP.

XX AAT63069;

DT 12-MAY-1997 (first entry)

DE DNA sequence of clone GLS-4.

XX Infectious bursal disease virus; IBDV; large segment; GLS 5 strain;

KW structural protein; VP2; VP3; VP4; immunogen; vaccine;

KW Gumboro disease; ss.

XX Infectious bursal disease virus.

Key Location/Qualifiers

FT CDS 1..1620

FT /\*tag= a

XX US5595912-A.

XX 21-JAN-1997.

XX 04-MAY-1990; 90US-0519202.

XX 23-MAR-1994; 94US-0216276.

XX 04-MAY-1990; 90US-0519202.

XX 28-JUN-1993; 93US-0083784.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Snyder D, Vakharia V;

XX WPI; 1997-107584/10.

XX P-PSDB; AAW14741.

XX Infectious bursal disease virus nucleic acids - for prodn. of

XX vaccines for poultry for protection against Gumboro disease

XX Claim 6; Column 65-70; 41pp; English.

XX The sequences given in AAT63066-69 represents fragments of the

XX infectious bursal disease virus (IBDV) large segment from the GLS 5

XX strain. These sequences represent clones GLS-1 to GLS-4. Proteins

XX derived from IBDV, esp. VP2 which is the major host protective immunogen

XX of IBDV, may be used to prepare vaccines against IBDV infection/Gumboro

XX disease.

XX Sequence 1622 BP; 420 A; 461 C; 425 G; 316 T; 0 other;

XX

XX

XX

XX

XX

alignment\_scores:

Quality: 77.00 Length: 125

Ratio: 1.305 Gaps: 6

Percent Similarity: 47.200 Percent Identity: 21.600

alignment\_block:

US-09-528-682-3 x AAT63069

XX

XX

XX

XX

Align seg 1/1 to: AAT63069 from: 1 to: 1622

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32

1 ATTCCACCAACAGATACCCAG.....CCAAT 29

32 uAspGlyArgTyrSerAsnPhelAlaLeuIleSerAlaGluGlyMetG 49

30 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAAGTGGTGC 73

49 lAspGlyAsp.....LeuPheGlyThr 56

74 AGGAAGGGGACAGATGTCATGTCGCAAGTGGGAGCCTAGCAGTGACG 123

57 ValAsnGlnSerAsnPhPro..... 63

124 ATTCAATGGTGGCAACTATCCAGGGGCCCTCCGTCACACCTAGTAGC 173

64 .....MetSerT 66

174 CTACGAAAGAGTGGCAACAGGATCTGCTTACGGTCGCTGGGGTGAGCA 223

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82

224 ACTTCGAGCTGTATCCCAATCTCTGAAGTACGAAAGAACCTGGTTACAGAA 273

83 AsnValLysTyrAsp.....MetAsnPhelLysLysLeuLeuArgPh 96

274 TAGGGCGATTTCACCCAGGAGCATGAACACAAATTGATA...CT 320

96 eMetGluAspPheIleGlyVal 104

321 GAGTGAGAGGGACCGCTTGGCATC 345

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAT05623

seq\_documentation\_block:

ID AAT05623 standard; cDNA; 3230 BP.

XX AAT05623;

DT 15-FEB-1996. (first entry)

DE IBDV GLS structural protein fragment cDNA.

XX IBDV; immunogen; vaccine; poultry; ss.

OS Infectious bursal disease virus of chickens.

Key Location/Qualifiers

FT CDS 114..3152

FT /\*tag= a

FT /product= VP2/VP4/VP3 structural protein fragment

XX WO9526196-A1.

XX 05-OCT-1995.

XX 29-MAR-1995; 95WO-US03772.

XX 29-MAR-1994; 94US-0219262.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Mengel-Whersat SA, Snyder DB, Vakharia V;

XX WPI; 1995-351197/45.

XX P-PSDB; AAR85439.

XX Chimeric infectious bursal disease virus polypeptide immunogens

XX comprising VP2 with modified epitopic amino acid sequence for

XX preparation of vaccines for immunisation of poultry

XX

XX

XX











OS Homo sapiens.  
 PN WO200155437-A2.  
 PD 02-AUG-2001.  
 PF 25-JAN-2001; 2001WO-US02623.  
 PR 25-JAN-2000; 2000US-0491404.  
 XX (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Drmanac RT;  
 DR WPI: 2001-451939/48.  
 DR P-PSDB; AAU14586.  
 XX  
 CC Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 PT  
 PS Claim 1; Page 776-777; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicit an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence encodes a protein of the invention.  
 XX  
 SQ Sequence 2663 BP; 510 A; 805 C; 667 G; 681 T; 0 other;

alignment\_scores:  
 Quality: 76.50 Length: 129  
 Ratio: 1.177 Gaps: 5  
 Percent Similarity: 50.388 Percent Identity: 26.357  
 alignment\_block:  
 US-09-528-682-3 x AAS22891/rev ..  
 Align seg 1/1 to reverse of: AAS22891 from: 1 to: 2663  
 1 AspPheThrArgAlaLeuGlnAlaTyCluProIleGluValAs 17  
 945 GAITTCCTTCATTCCTGGAGCATCGCTTTGGGCCAGTGATGCCAT 896  
 17 nThrAsnThrValThr.....GlnIleAsnGlySerAsnG 29  
 895 GGCGTGGCGCTATACTATTTTGAATAATACAGATCTTCCACTCCACG 846  
 29 luVal.....ProLeuAspGlyArgTyrSerAsnPhe 39  
 845 AGGTTTACGTCAGTCTCTATGTCAGTCTTGTGGAAGCGAGTGAGAA 796

40 AlaLeuIle..... 42  
 795 GTGTATGTCACCCAGAGGACAGACTAGCCAGCTGCTGAAGGAGATGAC 746  
 43 .....SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyT 56  
 745 AATGCTGACAGTTCAGACGAGGGCTACTAACCATGGAGCAGTTCACAA 696  
 56 hrValAsnGlnSerAsnProMetSerThrPheGluGlnValProAsn 72  
 695 CTGTCTCAAGAGTACCTTCCTCTCAAGACAGAGCAATCCAGGAG 646  
 73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAsp...Me 88  
 645 CTGATGGAGGAGGGGCTGGCATCCAGCAGCAGCAATGCAGACTTGCCT 596  
 88 tAsnPheLysLysLeuLeuArgPheMetGluAspAsp 100  
 595 CAACATCCGCTCACTG.....TTTATGGAGGATGAG 565  
 seq\_name: /SID55/gcgdata/geneseq/geneseq-nbml/NA1991.DAT:AAQ14897  
 seq\_documentation\_block:  
 ID AAQ14897 standard; DNA; 3228 BP.  
 AC AAQ14897;  
 DT 24-FEB-1992 (first entry)  
 XX GLS clones 1 to 4.  
 DE Infectious bursal disease virus; IBDV; vaccine; poultry; ss.  
 KW Infectious bursal disease virus.  
 OS Infectious bursal disease virus.  
 FH Key Location/Qualifiers  
 CDS 79..3150  
 FT /tag= a  
 FT /note= "Polypeptide encoded by GLS clones 1 to 4"  
 FT misc\_feature 1..348  
 FT /tag= b  
 FT /note= "GLS-1 clone"  
 FT misc\_feature 283..1252  
 FT /tag= c  
 FT /note= "GLS-2 clone"  
 FT misc\_feature 1722..3230  
 FT /tag= d  
 FT /note= "GLS-3 clone"  
 FT misc\_feature 999..2620  
 FT /tag= e  
 FT /note= "GLS-4 clone"  
 XX WO9116925-A.  
 PN 14-NOV-1991.  
 PD 30-APR-1991; 91WO-US03056.  
 PF 04-MAY-1990; 90US-0519202.  
 PR (UYMA-) UNIV MARYLAND COL.  
 PA Vakharia V;  
 PI WPI: 1991-353531/48.  
 DR P-PSDB; AAR15054.  
 XX Infectious bursal disease virus nucleic acids - useful as  
 PT vaccines against IBDV in poultry  
 PS Example; Page 30; 45pp; English.  
 XX











FD 11- MAY - 1954.

```
seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.AAQ31857
seq_documentation_block:
ID   AAQ31857 standard; DNA; 3115 BP.
```



```

XX AC AAQ31857;
XX DT 16-APR-1993 (first entry)
XX DE IBDV strain EDGAR segment A amplified by PCR.
XX KW Infectious Bursal Disease Virus; Gumboro's Disease; Birnaviridae;
XX KW chicken; Gallus; polymerase chain reaction; vaccine; ss.
XX OS Infectious Bursal Disease Virus.
XX FH
XX FT Key Location/Qualifiers
XX FT CDS 77..3115
XX FT /*tag= a
XX FT /product= 110kD-precursor
XX FT /note= "VP2, VP3 and VP4 are generated by cleavage
XX FT of the precursor encoded by ORF1"
XX FT CDS 43..480
XX FT /*tag= b
XX FT /label= ORF_2
XX FT /note= "function as yet unknown"
XX FT CDS 11..43
XX FT /*tag= c
XX FT /label= ORF_3
XX FT /note= "function as yet unknown"
XX FT primer_bind complement (19..60)
XX FT /*tag= d
XX FT /label= Primer_0_binding_site
XX FT primer_bind complement (50..81)
XX FT /*tag= e
XX FT /label= Primer_1_binding_site
XX FT primer_bind 574..604
XX FT /*tag= f
XX FT /label= Primer_1b_binding_site
XX FT primer_bind 1147..1178
XX FT /*tag= g
XX FT /label= Primer_2_binding_site
XX FT primer_bind complement (1179..1208)
XX FT /*tag= h
XX FT /label= Primer_3_binding_site
XX FT primer_bind 1817..1849
XX FT /*tag= i
XX FT /label= Primer_4_binding_site
XX FT primer_bind complement (1844..1883)
XX FT /*tag= j
XX FT /label= Primer_5_binding_site
XX PN EPS17292-A.
XX PD 09-DEC-1992.
XX PF 19-MAY-1992; 92EP-0201406.
XX PR 27-MAY-1991; 91BE-0000507.
XX PA (SOLV ) SOLVAY SA.
XX PI Colau D, De Wannemaeker C, Malarne D, Thiry G, Malarne D;
XX WPI; 1992-408761/50.
XX DR P-PSDB; AAR29277, AAR29278, AAR29279.
XX PT New recombinant avipox virus for use in vaccines - contg.
XX PT sequence encoding heterologous antigen in inter-gene region of
XX PT terminal inverted repeat
XX PS Example 19; Fig 9; 53pp; French.
XX CC IBDV EDGAR strain viral RNA was isolated from infected chickens and
XX CC reverse transcribed to cDNA using four pairs of primers. The primer
XX CC sequences were designed based on known IBDV segment A sequences from
XX CC the Australian strain 002-73, the German strain CU-1 and the British

```

```

CC strain 52/70 (see AAQ31858-Q31865 for sequences of primers 0, 1, 1b,
CC 2, 3, 4, 5 and 6, respectively; the primer 6 binding site is not
CC shown in the Features Table as this primer binds immediately
CC downstream of the segment A sequence). The resulting ds cDNA fragments
CC were subcloned, sequenced and aligned to give the full-length sequence.
CC The EDGAR strain sequences which were replaced by using primers based
CC on other sequences were then determined using primers situated outside
CC them and based on the EDGAR sequence. Segment A contains 3 ORFs, of
CC which ORF1 is the most important. The PCR-amplified segment A fragment
CC was inserted into a non-coding region of a transfer vector derived from
CC an attenuated Fowlpox virus, esp, into one of the terminal inverted
CC Repeat regions. The recombinant vector was then used to vaccinate chicks
XX against Gumboro's disease.
XX SQ Sequence 3115 BP; 843 A; 856 C; 799 G; 617 T; 0 other;

alignment_scores:
    Quality: 73.00      Length: 125
    Ratio: 1.259      Gaps: 6
    Percent Similarity: 46.400      Percent Identity: 21.600

alignment_block:
US-09-528-682-3 x AAQ31857 ..
Align seg 1/1 to: AAQ31857 from: 1 to: 3115

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
   ::: ||||| ::|||
962 ATTCCGACCAACGAGATAACCCAG.....CCAAAT 990

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetG 49
   : ::: : : : : : : : : : : : : : : : : : : : : : : : :
991 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAAGTGGCGGTC 1034

49 InAspGlyAsp.....LeuPheGlyThr 56
   || |||||
1035 AGCAGGGGACCAAGATGTCATGTGCGCAAGTGGAGCCTAGCAGTGACA 1084

57 ValAsnGlnSerAsnPhePro.....MetSert 66
   : : : : : : : : : : : : : : : : : : : : : : : : :
1085 ATCCATGGTGGCAACTATCCAGGGGCCCTCCGTCGCCTCACACTAGTAGC 1134

64 .....MetSert 66
   : : : : : : : : : : : : : : : : : : : : : : : : :
1135 CTACGAAGAGTGGCAACAGAGATCCCTGTTACGGTAGCGGGGTGAGCA 1184

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
   : ||||| : ||||| : : : : : : : : : : : : : : : : : : :
1185 ACTTCGAGCTGATCCCAATCTGAACTAGCAAGAACCTGGTTACAGAA 1234

83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
   : : : : : : : : : : : : : : : : : : : : : : : : :
1235 TAGCGCCGATTTGACCCAGGAGCATGACATACACAAATTTGATA...CT 1281

96 eMetGluAspPheIleGlyVal 104
   : ||||| : : : : : : : : : : : : : : : : : : :
1282 GACTGAGAGGACCGCTCTTGGCATC 1306

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ42757
seq_documentation_block:
ID AAQ42757 standard; DNA; 3257 BP.
XX
XX AC AAQ42757;
XX DT 25-NOV-1993 (first entry)
XX DE IBDV gene sequence.
XX KW Attenuated; vaccine; herpes virus; non-primate; live; safer; IBR;
XX KW infectious bovine rhinotracheitis; MDV; Marek's disease virus; fowl;
XX KW pseudo-rabies; swine; infectious bursal disease virus; ss.

```



```

XX Infectious bursal disease virus.
OS Key Location/Qualifiers
FH CDS 129..3167
FT /*tag= a
XX
XX US223424-A.
XX
XX 29-JUN-1993.
XX
XX 27-JUL-1988; 88US-0225032.
XX
XX 06-SEP-1985; 85US-0773430.
XX 27-JAN-1986; 86US-0823102.
XX 17-JUL-1986; 86US-0887140.
XX 02-SEP-1986; 86US-0902887.
XX 20-NOV-1986; 86US-0933107.
XX 27-JUL-1987; 87US-0078519.
XX 27-JUL-1988; 88US-0225032.
XX
XX (PRUT-) PRUTECH RES & DEV.
XX
XX Chiang CH, Cochran MD, Macdonald RD;
PI
XX WPI; 1993-219585/27.
XX
XX Recombinant fusion proteins for vaccine - comprises antigenic
PT sequences fused to viral sequences e.g. pseudo-rabies virus, used
PT as vaccines
XX
XX Disclosure; Fig 27; 127pp; English.
XX
XX The sequence is that of the infectious bursal disease virus (IBDV)
CC gene sequence from which may be used in the prodn. of attenuated
CC non-primate herpes viruses. These can be used as live vaccines and
CC provide a safer vaccine than currently available for e.g. pseudorabies
CC virus of swine, infectious bovine rhinotracheitis (IBR) virus or
CC Marek's disease of fowl.
XX
XX Sequence 3257 BP; 876 A; 906 C; 845 G; 630 T; 0 other;
SQ

```

alignment\_scores:

Quality:	73.00	Length:	125
Ratio:	1.259	Gaps:	6
Percent Similarity:	46.400	Percent Identity:	21.600

alignment\_block:

```

US-09-528-682-3 x AAQ42757 ..
Align seg 1/1 to: AAQ42757 from: 1 to: 3257
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
   :: ||||| ::|||::|
1014 ATTCCAAACAAACGAGATAACCCAG.....CCAAT 1042
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
   : :: ::|||::|
1043 CACATCC.....ATCAACTGGAGATAGTGACCTCCAAAAGTGGTC 1086
49 InAspGlyAsp.....LeuPheGlyThr 56
|| |||||
1087 AGCAGGGGATCACATGTTATGTTCGCGCAGAGGAGCCTAGCAGTGC 1136
|| |||||
57 ValAsnGlnSerAsnPhePro..... 63
::|::|
1137 ATCCATGTGTGGCAACTATCCAGGGGCCCTCGTCCCGTCAGCGTAGTGGC 1186
64 .....MetSerT 66
1187 CTACGAAGAGTGGCAACAGATCCGTCGTTACGTCGTCGGGTGAGCA 1236

```

```

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
   ::|||::| ::|||::|
1237 ACTTCGAGCTGATCCCAATCTCTGAACCTAGCAAGAACCTGGTTACAGAA 1286
83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
   ::|||::| ::|||::|
1287 TAGGGCCGATTTGACCCAGGAGCGCATGAACACACAAATTGATA...CT 1333
96 eMetGluAspAspPheIleGlyVal 104
   : ||| ||| ::|||::|
1334 GACTGAGAGGGACCGCTTGGCATC 1358
seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT31897
seq_documentation_block:
ID AAAT31897 standard; cDNA; 3257 BP.
XX
XX AC AAAT31897;
XX
XX DT 27-SEP-1996 (first entry)
XX
XX DE IBDV large cDNA fragment.
XX
XX KW IBDV; herpesvirus of turkey; vaccine; vector; S-HVT-003; ss.
XX
XX OS Infectious bursal disease virus of chickens strain S40747.
XX
XX FH Key Location/Qualifiers
XX CDS 129..3167
XX FT /*tag= a
XX
XX PN US5506128-A.
XX
XX PD 09-APR-1996.
XX
XX PF 06-SEP-1985; 85US-0773430.
XX
XX 27-JUL-1988; 88US-0225032.
XX 06-SEP-1985; 85US-0773430.
XX 27-JAN-1986; 86US-0823102.
XX 17-JUL-1986; 86US-0887140.
XX 02-SEP-1986; 86US-0902887.
XX 20-NOV-1986; 86US-0933107.
XX 27-JUL-1987; 87US-0078519.
XX 17-JUN-1993; 93US-0078873.
XX
XX (PRUT-) PRUTECH RES & DEV PARTNERSHIP.
XX
XX Chiang CH, Cochran MD, Macdonald RD;
XX
XX WPI; 1996-340868/34.
XX
XX Recombinant bovine rhinotracheitis virus expressing antigenic
PT polypeptide - useful as multivalent vaccine, esp. against
PT parainfluenza or rotavirus
XX
XX Example 21; Fig 27A-B; 114pp; English.
XX
XX The infectious bursal disease virus (IBDV) large DNA segment
CC (AAAT31897) contains 1 open reading frame that codes for 3 proteins,
CC 2 of which are antigens to provide protection against IBDV
CC infections of chickens. S-HVT-003 is a recombinant herpesvirus
CC of turkey (HVT) contg. the E. coli beta-galactosidase gene plus the
CC IBDV gene, both under the control of the pseudorabies virus gpX
CC gene promoter. Serocconversion of chickens to IBDV upon vaccination
CC with S-HVT-003 was demonstrated.
XX
XX Sequence 3257 BP; 875 A; 907 C; 845 G; 630 T; 0 other;
SQ

```

alignment\_scores:

Quality:	73.00	Length:	125
----------	-------	---------	-----



```

Ratio: 1.259          Gaps: 6
Percent Similarity: 46.400  Percent Identity: 21.600

alignment_block:
US-09-528-682-3 x AAT31897  ..

Align seg 1/1 to: AAT31897 from: 1 to: 3257

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
   :: ||||| ::||| |||
1014 ATTCCAACAACAGCAGATAACCCAG.....CCAAT 1042

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetG 49
   : :: ::::: ::::: ||||| |||
1043 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAAGGTGGTGC 1086

49 InAspGlyAsp.....LeuPheGlyThr 56
   || ||||| |||
1087 AGGCAGGGGATCAGATGTTATGTCGGCAAGAGGAGGCCTAGCAGTGACG 1136

57 ValAsnGlnSerAsnPhePro.....MetSert 66
   ::::: ::||| ::|||
1137 ATCCATGTGTGGCAACTATCCAGGGGCCCTCCGTCGCTCACGCTAGTGGC 1186

64 .....MetSert 66
   ::||| ::|||
1187 CTACGAAAGAGTGGCAACAGGATCGCTGTTACGGTCTGGGGTGAGCA 1236

66 hrPheGluGlnValProAsnLysGluPheLysGlyValIleSerAla 82
   ::||| ::||| ::||| ::||| ::||| ::|||
1237 ACTTCGAGCTGATCCCAANTCTGAACTAGCAAGAAGAACTGGTTACAGAA 1286

83 AsnValLysTyrAsp.....MetAsnPheLysLeuLeuArgPhe 96
   ::||| ::||| ::||| ::|||
1287 TACGCCGATTTGACCCAGGACCCATGAACATACACAAATTGATA...CT 1333

96 eMetGluAspPheIleGlyVal 104
   : ||| |||
1334 GAGTGAGAGGACCGCTCTTGGCATC 1358

```

seq\_name: /SIDS5/qcdata/qeneseq/qeneseqn-emb1/NA1998.DAT:AAV26605

seq_documentation_block:	
ID	AAV26605 standard; cDNA; 3261 BP.
XX	
XX	AAV26605;
XX	
DT	28-AUG-1998 (first entry)
XX	
DE	Plasmid pUC19FLAD78 containing IBDV segment A cDNA.
XX	
KW	Plasmid pUC19FLAD78; IBDV; Gumboro disease; vaccine;
KW	synthetic RNA transcript; reverse genetics; VP2; VP3; VP4; VP5;
KW	ss; cyclic; circular.
XX	
OS	Infectious bursal disease virus strain D78.
XX	
Key	Location/Qualifiers
CDS	97..534
FT	/*tag= a
FT	/product= VP5
FT	131..3169
CDS	
FT	/*tag= b
FT	/product= VP2-VP4-VP3
XX	
XX	WO9809646-A1.
XX	
PD	12-MAR-1998.
XX	
XX	31-JUL-1997; 97WO-US12955.
PF	
XX	05-SEP-1996; 96US-0708541.
PR	
XX	

PA	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
XX	
PI	Mundt E, Vakharia VN;
XX	
DR	WPI: 1998-193322/17.
DR	P-PSDB; AAW54374-75.
XX	
PT	Generation of live birnavirus from synthetic RNA transcripts -
PT	useful for vaccines against infectious bursal disease
XX	
PS	Example 1; Fig 5A-B; 84pp; English.
XX	
CC	Plasmid pUC19FLAD78 contains infectious bursal disease virus (IBDV)
CC	strain D78 segment A cDNA that encodes the polyprotein VP2-VP4-VP3
CC	(see AAW54375) and the newly identified non-structural protein VP5
CC	(see AAW54374). It was obtained by RT-PCR amplification (see
CC	AAW26592-95) of strain D78 RNA and cloning into vector pUC19. In a
CC	non reverse genetics system for IBDV, 3 independent full-length
CC	cDNA clones (see also AAW26604 and AAW26606) were constructed which
CC	contain segment A of serotype I strain D78 or serotype II strain
CC	23/82 and segment B of the serotype I strain P2. Synthetic RNAs of
CC	segments A and B were produced by in vitro transcription of
CC	linearised plasmids with T7 RNA polymerase. Transfection of Vero
CC	cells with combined plus strand transcripts of both segments
CC	generated infectious virus. The development of a reverse genetics
CC	system for double-stranded RNA viruses will facilitate studies of
CC	the regulation of viral gene expression pathogenesis, and design of
CC	a new generation of live and inactivated vaccines.
XX	
QQ	Sequence 3261 BP; 873 A; 909 C;847 G; 632 T; 0 other;

```
alignment_scores:
  Quality: 73.00      Length: 125
  Ratio: 1.259        Gaps: 6
  Percent Similarity: 46.400      Percent Identity: 21.600

alignment_block:
  US-09-528-682-3 x AAV26605 ..
  Align seq 1/1 to: AAV26605 from: 1 to: 3261
```

```

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
   ::: ||||| :::| |||||
1016 ATTCCAACAACAGAGATACCCAG.....CCAAAT 1044
   ::::
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
   ::::
1045 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAGTGGTGC 1088
   ::::
49 lnAspGlyAsp.....LeuPheGlyThr 56
   || ||||| |||
1089 AGCGAGGGGATCAGATGTCATGGTCGGCAAGAGGGGACCTAGCAGTGACG 1138
   ::::
57 ValAsnGlnSerAsnPhePro..... 63
   ::::: |||||: |||
1139 ATCCATGTTGGCAACTATCCAGGGGCCCTCGCTCCGTCAGCTAGTGGC 1188
   ::::
64 .....MetSerT 66
   ::::: |||||:
1189 CTACGAAAGAGTGGCAACAGGATCGTCTGTACGTCGCTGGGGTAGCA 1238
   ::::
66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
   ::|||:: |||||:: |||
1239 ACTTCGAGCTGTATCCCAAATCCTCAACTAGCAAGAAGACCTGGTTACAGAA 1288
   :::::
83 AsnValLysTyrAsp.....MetAsnPhelLysLeuLeuArgP 96
   ::::: |||||: |||||
1289 TAGCGCCGATTTGACCCAGGACCATGAACATPACACAAATTGATA...CT 1335
   ::::
96 eMetGluAspAspPheIleGlyVal 104
   :: ||| |||
1336 GAGTGAAGAGGACCGCTCTTGGCATC 1360

```



















































1-800-4-A-STATE

[illegible][illegible]







CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2937 BP; 1192 A; 480 C; 601 G; 664 T; 0 other;

alignment\_scores:  
 Quality: 70.00 Length: 117  
 Ratio: 1.129 Gaps: 5  
 Percent Similarity: 52.991 Percent Identity: 23.077

alignment\_block:  
 US-09-528-682-3 x AAS51819 ..

Align seg 1/1 to: AAS51819 from: 1 to: 2937

```

13 ProfileGluValAsn.....ThrAsnThrValThrGlnIleAsnG1 26
||||| ||| ::|||: :|||: |||
928 CCAATAGAATATAATCCAGGAGTAAGTAATAGCTATGCTAATGTAATGG 977

26 ySerAsnGluValProLeuAspGly.....ArgTyrSerAsnPheAlaL 41
||||| |||: ||| |||: |||: |||
978 TTCATTGAACCTTCGATAAGGAATAATAGATTCACTCATGTGGCTT 1027

41 euLeSerAlaGluGlyMetGlnAspGly.....AspLeuPheGly 55
||||| :|||: ||| :|||: |||
1028 ACATTAAGCCACAAATGGGCATAATCAGATAGCGTTTCAATTACTGGT 1077

56 ThrValAsnGln..... 59
|||||: |||

1078 ACATTAACCAAGGTAGTAAAGCAGATGGGAGGCTCCTACTGTAAAGT 1127

60 .....SerAsnPheProMetSerThrPheGluG 69
|||||: |||: |||: |||: |||: |||
1128 ATATGAAGTTTAAAGGATGCTAATGAATACCAAAAGTGTATACGCAA 1177

69 InValProAsnAsnLysGluPheLysGlyVal..... 79
||||| :|||: ||| |||
1178 ATGTATCAGATCTTCATGTTTAAAGATGTAAGTGAAGAAATGAAGAT 1227

80 .....IleSerAlaAsnValLysTyrAspMetAsnPheLysLysLe 93
||||| ||| :||| ||| :|||: |||
1228 AAATTAAGGTAGAAAATAATGGGAAGTATAAATAGATATTGAAATTT 1277

93 u 93
1278 A 1278

```

seq\_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAS54951

```

seq_documentation_block:
ID AAS54951 standard; DNA; 3006 BP.
XX
AC AAS54951;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1263.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.

```

```

PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR P-PSDB; AAU37092.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
PS Claim 27; Seq ID No 8588; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3006 BP; 1220 A; 490 C; 616 G; 680 T; 0 other;

```

```

alignment_scores:
  Quality: 70.00 Length: 117
  Ratio: 1.129 Gaps: 5
  Percent Similarity: 52.991 Percent Identity: 23.077

alignment_block:
US-09-528-682-3 x AAS54951 ..
Align seg 1/1 to: AAS54951 from: 1 to: 3006

13 ProfileGluValAsn.....ThrAsnThrValThrGlnIleAsnG1 26
||||| ||| ::|||: :|||: |||
997 CCAATAGAATATAATCCAGGAGTAAGTAATAGCTATGCTAATGTAATGG 1046

26 ySerAsnGluValProLeuAspGly.....ArgTyrSerAsnPheAlaL 41
||||| |||: ||| |||: |||: |||
1047 TTCATTGAACCTTCGATAAGGAATAATAGATTCACTCATGTGGCTT 1096

41 euLeSerAlaGluGlyMetGlnAspGly.....AspLeuPheGly 55
||||| :|||: ||| :|||: |||
1097 ACATTAAGCCACAAATGGGCATAAATCAGATAGCGTTTCAATTACTGGT 1146

56 ThrValAsnGln..... 59
|||||: |||

1147 ACATTAACCAAGGTAGTAAAGCAGATGGGAGGCTCCTACTGTAAAGT 1196

60 .....SerAsnPheProMetSerThrPheGluG 69
|||||: |||: |||: |||: |||: |||
1197 ATATGAAGTTTAAAGGATGCTAATGAATACCAAAAGTGTATACGCAA 1246

69 InValProAsnAsnLysGluPheLysGlyVal..... 79

```



```

:::||||| :||||| :||||| :|||
1247 ATGTATCAGATTCTTCAAGTCTTAAGATGTAAGTCTGAGAAATGAAAGAT 1296
80 .....IleSerAlaAsnValLysTyrAspMetAsnPheLysLysLe 93
1297 AATTTAAAGTAGAATAATATGGAAGTATATAATATAGATATTGAAAAATT 1346
93 u 93
|
1347 A 1347

```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH41061

seq\_documentation\_block:  
ID AAH41061 standard; RNA; 3148 BP.

AC AAH41061;

DT 28-AUG-2001 (first entry)

DE Infectious bursal disease virus genomic RNA fragment.

KW Infectious bursal disease; IBD; chicken; turkey; strain YH-91-CLC;  
KW vaccine; GS.

OS Infectious bursal disease virus of chickens.

FH Key Location/Qualifiers

FT CDS 39..1394

FT /tag= a

FT /partial

FT /product= "IBDV protein"

FT CDS 1395..2204

FT /tag= b

FT /partial

FT /product= "IBDV protein"

FT CDS 2205..3074

FT /tag= c

FT /partial

FT /product= "IBDV protein"

XX JP2001086983-A.

XX 03-APR-2001.

XX 21-JUL-1999; 99JP-0206771.

XX 21-JUL-1998; 98JP-0204615.

PR 21-JUL-1999; 99JP-0205569.

XX (EISA ) EISAI CO LTD.

XX WPI; 2001-338419/36.

DR P-PSDB; AAG62331, AAG62333, AAG62334.

XX Infectious bursal disease virus YH-91-CLC strain A, useful as a vaccine

PT for poultry -

XX Claim 4; Page 10-14; 23pp; Japanese.

XX An infectious bursal disease (IBD) virus of chickens YH-91-CLC strain can

CC be used as the active agent of a an IBD vaccine. The present sequence

CC represents a YH-91-CLC strain genomic RNA fragment which encodes peptides

CC which can be used in the vaccine. The vaccine can be used to treat IBD

CC infected chickens and turkeys.

XX Sequence 3148 BP; 842 A; 883 C; 818 G; 605 U; 0 other;

alignment\_scores:

Quality: 70.00 Length: 127

Ratio: 1.186 Gaps: 6

Percent Similarity: 46.457 Percent Identity: 21.260

alignment\_block:

US-09-528-682-3 x AAH41061

Align seg 1/1 to: AAH41061 from: 1 to: 3148

14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30

918 AUUGUGAUUCCCAACCGCAGAGUAACCCAG..... 947

30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47

948 .CCAACACAUCC.....AUCAAACUGGAGAUAGUGACCUCCAAAAGUG 990

47 lYMetGlnAspGlyAsp.....LeuPhe 54

991 GUGUGACGGGGGGAUACAUGUAGUUGGUCAGCAAGUGGGAGCCUAGCA 1040

55 GlyThrValAsnGlnSerAsnPhePro..... 63

1041 GUGACGAUCCACGGGCGCAACUACCCAGGGGCCUCCUCCGUCACACU 1090

64 .....M 64

1091 AGUAGCCUACGAAAGAGUGGCAACAGGAUCUGUCGUACGCGCGGG 1140

64 etSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIle 80

1141 UGAGCAACUUCGAGCUGAUCCCAAAACCCUGAACUAGCAAAAGAACCCUGUC 1190

81 SerAlaAsnValLysTyrAsp.....MetAsnPheLysLysLeuLe 94

1191 ACAGAAUACGCCCGAUUUGACCCAGGGGCCCAUGAACUACACACAAAUAUGAU 1240

94 uArgPheMetGluAspPheIleGlyVal 104

1241 A...CUGAGUGAGAGGAGCGGUCUUGGCAUC 1268

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF25061

seq\_documentation\_block:

ID AAF25061 standard; cDNA; 3260 BP.

XX AAF25061;

XX 30-APR-2001 (first entry)

XX Nucleotide sequence of the A-segment of an IBDV isolate.

DE IBDV; vaccine; infectious bursal disease; IBD; poultry; chicken; duck;

KW rooster; hen; broiler; roaster; breeder; layer; turkey; ss.

XX Infectious bursal disease virus.

OS EP1069187-A1.

XX 17-JAN-2001.

XX 14-JUL-1999; 99EP-0202316.

XX 14-JUL-1999; 99EP-0202316.

XX (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.

XX WPI; 2001-184296/19.

XX New infectious recombinant Infectious Bursal Disease Virus, useful as a

PT vaccine for preventing subclinical and clinical infectious bursal

PT disease in poultry, e.g. chickens, broilers, roasters, breeders,

PT layers, turkeys or ducks -

XX Disclosure; Page 41-42; 90pp; English.

XX







```

54  etSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIle 80
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
1234  TGAGCAACTTCGAGCTGATCCCAATCTCTGAAC TAGCAAGAACCTGGTC 1283
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::

81  SerAlaAsnValLysTyrAsp.....MetAsnPheLysLysLeu 94
    ::::|||::|||::|||::|||::|||::|||::|||::|||::
1284  ACAGAATAAGGCGGATTTGACCCAGGGGCCATGAAC TACACAAATTTGAT 1333
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::

94  uArgPheMetGluAspAspPheIleGlyVal 104
    ::::|||::|||::|||::|||::|||::|||::|||::|||::
1334  A...CTGAGTGTAGAGGAGCCGCTCTTTGGCATT 1361
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::

```

seq\_name: /SIDS5/qcdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ60605

```
seq_documentation_block;
ID AAZ60605 standard; DNA; 5398 BP.
```

AAZ60605;

DT 16-MAY-2000 (first entry)

DE Nucleotide sequence encoding a yeast Sos-related protein.

Ras; Son of sevenless; Sos; crystal; Ras-Sos complex; cancer;  
KW  
KW  
protein coordinate data: ss

OS *Saccharomyces cerevisiae*.

PN WO200005258-A1.

PD 03-FEB-2000.

PF 20-JUL-1999; 99WO-US16348.

PR 21-JUL-1998; 98US-0119794.

PA (UYRQ ) UNIV ROCKEFELLER.

PI Borlack-Sjodin A, Marqarit SM, Bar-Saqi D, Cole P, Kuriyan J;

DR WPI; 2000-182647/16.

Novel crystals comprising a Ras-Son of sevenless complex, useful for screening drugs useful in cancer treatment -

PS Disclosure; Page 201-203; 224pp; English;

The specification describes a crystal complex comprising at least a Ras contacting region of a Ras protein and at least a Son of sevenless (Sos) protein catalytic region fragment, that effectively diffracts x-rays. Ras and Sos form a tight complex. Sos does not impede the binding sites for the nucleotide base and the ribose of GTP or GDP and thus the Ras-Sos complex maintains a structure that permits nucleotide release and rebinding. The crystals are used for the determination of the atomic coordinates of the complex to a resolution of more than 5.0 Angstrom. The crystals, or a dataset comprising the three-dimensional coordinates obtained from the crystals, is useful for identifying an agent that stabilizes the Ras-Sos complex. The crystals are also useful for identifying agents that inhibit the formation of Ras-Sos complex. Ras and Sos fragments are useful for growing a crystal of a protein-ligand complex. Agents that stabilize or inhibit the formation of Ras-Sos complex are useful in the treatment of cancer. The present sequence encodes a Sos-related protein.

Sequence 5398 BP: 1849 A: 984 C: 985 G: 1580 T: 0 other: 22

alignment_scores:		
Quality:	70.00	Length: 101
Ratio:	1.250	Gaps: 4
Percent Similarity:	55.446	Percent Identity: 25.743

```

alignment_block:
  US-09-528-682-3 x AAZ60605      ..
  Align seg 1/1 to: AAZ60605 from: 1 to: 5398

  21 ValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArg..... 35
  ||| :::::::::: ::::::::::::::::::::|||:::
  2044 GTAGTAGTGTCATGATCTAGTTTCAGTTCCTGTCGGTAAACGTAA 2093

  36 .....TyrSerAsnPheA 40
  2094 TGTAAGTACCCAGTACAACGGATACATTGACTCCAATGAGATCATCATCA 2143
  40 IaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThr 56
  :: ::::: |||::: ::::::::::::::|
  2144 GTACAGCTCAAT...GAGACGATATGGAAAAATTTCTCAGCTCTAGGTCCA 2190

  57 ValAsnGlnSerAsn.....PheProMetSe 65
  |||::: ||| ::::
  2191 AGAAATAGTGGTTAAATTCGTCTGTACACCAAGGACTTCAATACAAATTC 2240

  65 rThrPheGluGlnVal...ProAsnAsnLysGluPheLysGlyValIles 81
  |||||::: |||::: |||::: |||::: |||::: |
  2241 TACTTTGGGAAGATTTTTCACCGCTCCAACAAAAATTTTAAAGTCAGCTAAAT 2290

  81 erAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPheMet 97
  || |||::: |||::: |||::: |||::: |||::: |||:::
  2291 CGATTACGAATGGTTGATGGGAATTCGGAATTTTAAAGGCATGTT 2340

  98 Glu 98
  ::
  2341 CAG 2343

```

seq\_name: /SID55/acadata/qeneseq/qeneseq-emb1/NA2000.DAT:AAA81463

seq\_documentation\_block:

XX  
 ITD C4470544  
 SECURITY, DNA, 70043 DE

AC  
XX  
AAAB1403;DT 04-DEC-2000 (first entry)  
yy

DE N. meningitidis partial DNA sequence gnm\_11 SEQ ID NO:11.

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW *Meningococcus* B; MenB; ds.

OS *Neisseria meningitidis*.

AA  
PN  
WO200022430-A2

XX  
PD  
20-APR-2000

XX  
PF 08-OCT-1999.  
99WU-UJ23573

00-0000-1000. 00000-0103701

PR 30-APR-1999; 99US-0132068.

PA (CHIR ) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;  
PI Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappuoli R, Pizza M;

DR WPT: 2000-318079/27.

Isolated nucleotide sequences of *Neisseria meningitidis* which can be used in the diagnosis and treatment of *N. meningitidis* infection and other *Neisseria* infections, for example, *N. gonorrhoea*.

XX PS Claim 7; Page 330-353; 1760pp; English.



CC The present invention describes methods of obtaining immunogenic  
CC proteins from *Neisseria* genomic sequences. AAA81453 to AAA82414  
CC represent specifically claimed *Neisseria meningitidis* genomic DNA  
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent  
CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to  
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the  
CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81322 to  
CC AAA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF  
CC sequences, which are all used in the exemplification of the present  
CC invention. The nucleic acid sequences, protein sequences, and antibodies  
CC against them, can be used in the manufacture of a composition. The  
CC composition can be used as a medicament (or in the manufacture of a  
CC medicament) for treating, preventing or diagnosing infection due to  
CC *Neisseria* bacteria. For example, some of the identified proteins could  
CC be components of vaccines against *Meningococcus* B; against all serotypes;  
CC and/or against all pathogenic *Neisseriae*. Identification of sequences  
CC from the bacterium will also facilitate production of biological probes,  
CC particularly organism-specific probes. Attempts to make efficacious  
CC *Meningococcus* B vaccines have failed mainly due to antigen tolerance.  
CC Multivalent vaccines have also been tried but none have successfully  
CC overcome antigenic variability. The provision of further, complete  
CC sequences may provide an opportunity to identify secreted or surface  
CC exposed proteins that may be presumed targets for the immune system and  
CC which are not antigenically variable or at least more conserved than  
CC other more variable regions.  
XX  
SQ Sequence 78845 BP; 19372 A; 21111 C; 19042 G; 19318 T; 2 other;

alignment\_scores:  
Quality: 70.00 Length: 123  
Ratio: 1.111 Gaps: 6  
Percent Similarity: 51.220 Percent Identity: 26.829

alignment\_block:

US-09-528-682-3 x AAA81463 ..

Align seg 1/1 to: AAA81463 from: 1 to: 78845

```
9 GlnAlaTyrGluProIleGluValAsnThrAsnThrValThrGlnIleAs 25
 8732 AAAGCCATAGATCCGATTTGCTTCCCTCACCGCGCTACCGCATCAA 8781
25 nGlySerAsnGluValProLeuAspGlyArgTyr.....SerAsnPheA 40
 8782 CGCGACGAAACCCCTTTCAGAGCGCATCCGTACAAATCGGGCACTACC 8831
40 IaLeu.....IleSerAla 44
 8832 TGCTGCTCAACCCCAACATCTCATTTGCTGCCAAGCAGACAGTCGCTG 8881
45 GluGlyGlyMetGln.....AspGlyAs 52
 8882 ACCGGAGGATCCCAATGCTGGCAGGAGCCGCCAGCGACGGACGGCAA 8931
52 pLeuPheGlyThrValAsnGlnSerAsnPhe.ProMetSerThrPheGlu 68
 8932 ACGGGAATCCTCCAGAAACATCCACTACGCCCATTTCGCGCAGGTT 8981
69 GlnVal.....ProAsnAsnLysGluPh 76
 8982 TCGGTTTCCACCAAAACACCGGCTTTAAACGCATCGCGACGTTTCAACGTT 9031
76 eIysGly.....ValIleSerAlaAsnValLysTyrAspMetAsnPheL 91
 9032 TCAGGGCAACAGCTCCGAACCTGAAATTTGGCGTACAGCATACATTTTA 9081
91 ysLysLeuLeuArgPhe 96
 9082 AGCAGGTTTGTGATTC 9098
```

seq\_name: /STD55/gcdata/genesec/genesecn-emb1/NA2000.DAT:AAF21608

seq\_documentation\_block:

ID AAF21608 standard; DNA; 349980 BP.

XX AAF21608;

DT 13-MAR-2001 (first entry)

DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO:109.

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;  
diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
ds.

OS *Neisseria meningitidis*.

PN WO2000066791-A1.

PD 09-NOV-2000.

PF 08-MAR-2000; 2000WO-US05928.

PR 30-APR-1999; 99US-0132068.

PR 08-OCT-1999; 99WO-US23573.

PR 28-FEB-2000; 2000GB-0004695.

XX (CHIR ) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;  
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V, Rappuoli R;  
PI Frazer CM, Grandi G;

XX WPI; 2000-647603/62.

XX *Neisseria meningitidis* B full length genome sequence and open reading  
frames are used to detect, treat and prevent *Neisseria* infections -

Claim 7; Appendix A; 692pp; English.

XX The present invention describes the full length genome of  
CC *Neisseria meningitidis* B (NMB). The sequences in AAF21544 and AAF21607  
CC to AAF21613 represent fragments of the NMB genomic sequence, as the  
CC sequence was too long to go in a record on its own it was split into 8  
CC sequences which overlap each other at the beginning and end of each  
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at  
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at  
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the  
CC *Neisseria* proteins given in AAB58550 to AAB58593, and AAF21589 to  
CC AAF21606 represent PCR primers which are used in the exemplification of  
CC the present invention. The NMB genome and fragments from it have  
CC antibacterial activity, and can be used in vaccines and gene therapy.  
CC *Neisseria* nucleic acids, proteins and/or antibodies which binds to the  
CC proteins can be used in compositions for treating or preventing infection  
CC due to *Neisseria* bacteria or as a diagnostic reagent for detecting the  
CC presence of *Neisseria* bacteria or of antibodies raised to *Neisseria*  
CC bacteria. Computers, computer memory, computer storage medium or computer  
CC databases can be used in a search to identify open reading frames (ORFs)  
CC or coding sequences within the NMB genome. The DNA sequences provide  
CC further opportunities to find antigenic or immunogenic proteins which are  
CC more effective in vaccines than the outer membrane proteins currently  
CC used.

XX Sequence 349980 BP; 82523 A; 82940 C; 96712 G; 87805 T; 0 other;

alignment\_scores:

Quality: 70.00 Length: 123  
Ratio: 1.111 Gaps: 6  
Percent Similarity: 51.220 Percent Identity: 26.829

alignment\_block:

US-09-528-682-3 x AAF21608/rev ..







```
seq_documentation_block:
ID      AAI99683 standard; DNA; 4403765 BP.
XX
XX      AAI99683;
AC
XX
XX      15-JAN-2002 (first entry)
DT
XX
XX      Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
DE
XX
XX      Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW      variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS      Mycobacterium tuberculosis.
XX
XX      US6294328-B1.
PN
XX
XX      25-SEP-2001..
PD
XX
XX      24-JUN-1998;    98US-0103840.
PF
XX
XX      24-JUN-1998;    98US-0103840.
PR
XX
XX      (GENO-) INST GENOMIC RES.
PA
XX
XX      Fleischmann RD, White OR, Fraser CM, Venter JC;
PI
XX      WPI; 2001-647261/74.
DR
XX
XX      Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT      determining the nucleotide sequence of the strain at positions in the
PT      genome corresponding to positions where M. tuberculosis strains CDC
PT      1551 and H37Rv differ -
XX
XX      Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
PS
XX
XX      The invention relates to evaluating strain variation within and between
CC      different populations of the tuberculosis bacterial pathogen, and
CC      Mycobacterium tuberculosis or related Mycobacterium by determining the
CC      nucleotide sequence of the first strain at positions in the complete
CC      sequence of the genome that correspond to positions that differ in the
CC      nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC      H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC      M. tuberculosis and has valuable application in the fields of
CC      tuberculosis genetics, epidemiology, patient treatment and epidemic
CC      monitoring.
CC
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from USPTO
CC      at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
XX      Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
SQ

alignment_scores:
Quality:    70.00          Length:    75
Ratio:      1.489         Gaps:      4
Percent Similarity: 62.667 Percent Identity: 28.000

alignment_block:
US-09-528-682-3 x AAI99683/rev ..

Align seg 1/1 to reverse of: AAI99683 from: 1 to: 4403765

          9   GlnAlatYrGluProlLeIGluValAsnThrAsnThrValThrGlnIleas 25
          4044858CGGCGGTACAGCCGGTTGATACCCTGGTG...GCGATAACACAGATCAA 4044812

          25 nGlySerAsn.....GluValProLeuaspGlyArgTyrSerAsnPheA 40
          4044811LCGGGCACACACGTCAAGACGCCACAGATCGACGATCCGCGACACAGCTGC 4044762

          40 laLeulleSerAlaGlUGlYGlyMetGln....AspGly..... 51
          ||||| ::::||||: :: |
```

404761GCTGATCCTGAAGAAGGTAAACAAGACCGACCACGGTCCGGANGGTC 4044712

52 .....AspLeuPheClyThrValas 58  
|||||||:::~::~:

4047411ATCACCAAGTACCCACCGGTACGCCGTGCACCTGTTCAACGCGCTCAG 4044662

58 nGlnSerAsnProMetSerThr 66  
: : ::||| : ::|||  
4044661CGCCAAAACCGAAGTCAGCACG 4044637

seq\_name: /SID5/gcgdata/geneseq/geneseqp-emb1/NA2001A.DAT:AAI199682

seq\_documentation\_block:  
ID AAI199682 standard; DNA; 4411529 BP.

XX AC XX  
XX AC XX  
DT XX  
XX  
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.  
XX  
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;  
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN US6294328-BL.  
XX  
PD 25-SEP-2001.  
XX  
PF 24-JUN-1998; 98US-0103840.  
XX  
PR 24-JUN-1998; 98US-0103840.  
XX  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Fleischmann RD, White OR, Fraser CM, Venter JC;  
DR WPI; 2001-647261/74.  
XX  
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises  
PT determining the nucleotide sequence of the strain at positions in the  
PT genome corresponding to positions where M. tuberculosis strains CDC  
PT 1551 and H37Rv differ  
XX  
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.  
XX  
CC The invention relates to evaluating strain variation within and between  
CC different populations of the tuberculosis bacterial pathogen,  
CC Mycobacterium tuberculosis or related Mycobacterium by determining the  
CC nucleotide sequence of the first strain at positions in the complete  
CC sequence of the genome that correspond to positions that differ in the  
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI199683) and  
CC H37Rv (AAI199682). The method is useful for evaluating strain variation of  
CC M. tuberculosis and has valuable application in the fields of  
CC tuberculosis genetics, epidemiology, patient treatment and epidemic  
CC monitoring.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=6294328BL.  
XX  
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

```

alignment_scores:
  Quality: 70.00      Length: 75
  Ratio: 1.489       Gaps: 4
  Percent Similarity: 62.667  Percent Identity: 28.000

alignment_block:
  US-09-528-682-3 x AA199682/rev ..

Align seq 1/1 to reverse of: AA199682 from: 1 to: 4

```











```

92 .....LysLeuLeuArgPheMetGlu 98
1025 CAGTACAAATAAGTAGTACTTCGCTTATATGAA 1065

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ14896

seq_documentation_block:
ID  AAQ14896 standard; DNA; 1471 BP.
XX
AC  AAQ14896;
XX
DT  24-FEB-1992 (first entry)
XX
DE  E/DEL-2 clone.
XX
KW  Infectious bursal disease virus; IBDV; vaccine; poultry; ss.
XX
OS  Chicken infectious bursal disease virus.
XX
FH  Location/Qualifiers
FT  19..1470
FT  /*tag= a
FT  /note= "polypeptide encoded by E/DEL-2 clone"
XX
PN  WO9116925-A.
XX
PD  14-NOV-1991.
XX
PF  30-APR-1991; 91WO-US03056.
XX
PR  04-MAY-1990; 90US-0519202.
XX
PA  (UYMA-) UNIV MARYLAND COL.
XX
PI  Vakharina V;
XX
DR  WPI; 1991-353531/48.
XX
DR  P-PSDB; AAR15053.
XX
PT  Infectious bursal disease virus nucleic acids - useful as
PS  vaccines against IBDV in poultry
XX
PS  Example; Page 28; 45pp; English.
XX
CC  The sequence is that of a clone of variant Delaware type E of
CC  infectious bursal disease virus (IBDV). It can be used in broad
CC  spectrum IBDV poultry vaccines, administration is ophthalmically,
CC  by injection, nasally or orally, at any time after hatching and as a
CC  booster at other times. Poultry treated include chickens, roosters,
CC  broilers, roasters, breeders, layers, turkeys and ducks. See also
XX  AAQ14897.
XX
SQ  Sequence 1471 BP; 403 A; 389 C; 366 G; 313 T; 0 other;

```

```

alignment_scores:
  Quality: 69.00      Length: 121
  Ratio: 1.211      Gaps: 5
  Percent Similarity: 47.107  Percent Identity: 23.140

alignment_block:
US-09-528-682-3 x AAQ14896 ..
Align seg 1/1 to: AAQ14896 from: 1 to: 1471

16 ValAsnThrAsnThrValThrGlnLeuAsnGlySerAsnGluValProLe 32
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
937 ATTCCAACCAATGAGTAGATACCCAGCAATCATCATCAACTGGAGAT 986
32 u.....AspGlyArgTyrSerAsnPheAlaLeuIleSerA 44
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
987 AGTGACCTCCAAAGATGTGTCAGGAGGGGACAGATGTCATGTCGG 1036

```

```

44 laGluGlyGlyMetClnAspGlyAspLeuPheGlyThrValAsnGlnSer 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1037 CA.....AGTGGAGCCTAGCAGTACGATCCATGGTGGC 1071

61 AsnPhePro.....
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1072 AACTATCCAGGAGCCCTCCGTCCTCCGTCACACTAGTGGCTACGAAAGAGT 1121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 .....MetSerThrPheGluGlnV 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1122 GGCAACAGGATCTGTGTTACGGTCGCTGGGGTGAGCAACTTCGAGCTGA 1171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70 alProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLysTyr 86
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1172 TCCCAATCCCTGAAGTACCAAGAACCTGGTTACAGAAATATGCGCGATTT 1221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
87 Asp.....MetAsnPheLysLysLeuLeuArgPheMetGluAspAs 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1222 CACCCAGGAGCCATCACTACACGAAATTGATA...CTGAGTGAGAGGA 1268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 pPheIleGlyVal 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1269 CCGGTTGGCATC 1281

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:AA60874

seq_documentation_block:
ID  AA60874 standard; DNA; 3129 BP.
XX
AC  AA60874;
XX
DT  29-OCT-1991 (first entry)
XX
DE  Large RNA segment of infectious bursal disease virus.
XX
KW  IBDV; poultry; vaccine.
XX
OS  Avian infectious bursal disease virus.
XX
FH  Location/Qualifiers
FT  2..3067
FT  /*tag= a
FT  /note= "Longest open reading frame"
XX
PN  WO8607060-A.
XX
PD  04-DEC-1986.
XX
PF  30-MAY-1986; 86WO-AU00156.
XX
PR  23-AUG-1985; 85AU-0002118.
PR  30-MAY-1985; 85AU-0000815.
PR  01-JAN-1986; 86AU-0059668.
XX
PA  (CSIR ) COMMONWEALTH SCIENT ORG.
XX
PI  Azad AA, Hudson PJ, Fahey KJ;
XX
DR  WPI; 1986-332075/50.
DR  P-PSDB; AAP61316.
XX
PT  Recombinant DNA corresp. to infectious bursal disease virus-RNA -
PT  used for producing polypeptide(s) for use in treatment of
PT  infectious bursal disease virus in chickens.
XX
PS  Claim 10; Fig 10; 78pp; English.
XX
CC  The avian IBDV 32KD structural protein product may be used in
CC  serological diagnosis and development of a vaccine against the
CC  disease. The protein is encoded by the only extensive reading frame
CC  of the sequence.
XX

```



SQ Sequence 3129 BP; 845 A; 882 C; 799 G; 603 T; 0 other;

## alignment\_scores:

Quality: 69.00 Length: 128  
Ratio: 1.211 Gaps: 5  
Percent Similarity: 44.531 Percent Identity: 21.094

## alignment\_block:

US-09-528-682-3 x AAN60874 ..

Align seg 1/1 to: AAN60874 from: 1 to: 3129

```
13 ProIleGluValAsnThrValThrGlnIleAsnGlySerAsnG1 29
   |||   ::   |||:::  |||||:::||||
917 CCAACCAAGTGAGTAACCCAGCCAGTTACATCCATT..... 952

29 uValProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluG 46
   :::::  :::::  :::::  :::::  :::::  :::::
953 .....AACTGGAGATAGTAACCTCCAAAA 977

46 LyGlyMetGlnAspGlyAsp.....Leu 53
   ::|||  |||  |||||
978 GTGGAGGTCAGGCTGGAGATCAGATGCTCTGTTGGCAAGTGGGAACCTA 1027

54 PheGlyThrValAsnGlnSerAsnPhePro..... 63
   |||:::  :::  |||:::
1028 GCAGTGACAATTCATGGTGGAAACTACCCAGGTGCCCTCCGCCCGTCAC 1077

63 ..... 63

1078 ACTAGTAGCCTATGAAGAGTGGCAACAGGATCTGTGTACGGTCGCTG 1127

64 ..MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1128 GGGTGAGCAACTTCAGCTGATCCCGCAATCCTGAGCTAGCCCAAGAACCTA 1177

80 IleSerAlaAsnValIlyTyrAsp.....MetAsnPheLysLysLe 93
   :::::  :::::  |||
1178 GTCACAGAAATATGCCGATTTGCCCCAGGAGCCATGAATACACAAACT 1227

93 uLeuArgPheMetGluAspPheIleGlyVal 104
   |:::  :::  |||  |||  :::|:::
1228 AATC...CTGAGTGAGAGGACCGCTTGGTATC 1258
```



OM of: US-09-528-682-3 to: EST:\* out\_format : pfs

Date: Jun 18, 2002 6:26 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO\_spo/US09528682/runat\_18062002\_082443\_7797/app\_query.fasta\_1.689  
-DB=EST -OEMT=fastap -SUFFIX=p2n.rst -GAPOP=12,000 -GAPEXT=4,000  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4,500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6,000  
-FGAPEXT=7,000 -XGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6,000  
-DELEXT=7,000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=1000 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=50 -MODE=LOCAL -OUTFWT=pfs -NORM=ext -HEAPSIZ=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09528682\_@CGN1\_1\_5309  
-NCPUP=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-3

Query length: 107

Database: EST:\*

Database sequences: 13736207

Database length: 1841457050

Search time (sec): 4056.470000

score\_list:

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
gb_est2:Bj030418	+	79.00	175.34	1.61	621	Bj030418 BJ030418 NIBB Moch11
gb_gss:BH298334	+	76.50	169.69	3.32	586	BH298334 CH230-1012.TJ CHORI-23
gb_gss:BH324419	+	76.50	168.64	3.80	660	BH324419 CH230-189M18.TV CHORI-23
gb_est2:BH316490	+	76.00	169.49	3.41	540	BH316490 saf03H08.y1 Gm-cl085 G
gb_est2:BM270616	+	76.00	169.23	3.52	554	BM270616 sak15011.y1 Gm-cl075 G
gb_est1:AL636005	+	76.00	166.48	5.01	723	AL636005 AL636005 XGC-neurula
gb_gss:BH157168	+	76.00	164.12	6.78	909	BH157168 ENTL34TR Entamoeba h
gb_gss:BH266151	+	75.00	172.40	2.35	322	BH266151 CH230-55LA.TV CHORI-23
gb_est1:AIL134083	+	75.00	166.28	5.15	583	AIL134083 GH11413.Sprime GH bro
gb_gss:AZ86197	+	74.50	164.15	6.76	637	AZ86197 2M0268L1GF Mouse 10kb
gb_est1:AA917243	+	74.00	166.05	5.30	471	AA917243 SMOVL3CAN1IG04 Onchoce
gb_est1:AA333548	+	73.50	169.69	3.32	294	AA333548 EST7697 Embryo, 8 wee
gb_gss:BH413108	+	73.50	163.27	7.56	548	BH413108 CH230-101K19.TV CHORI-
gb_est1:AUI18471	+	73.50	158.72	13.56	852	AUI18471 AUI18471 HEMBA1 Homo
gb_est2:BH316465	+	73.00	162.92	7.91	504	BH316465 saf03F02.y1 Gm-cl085 G
gb_est2:BE590377	+	73.00	161.74	9.20	565	BE590377 SB117 Sugar Beet germi
gb_est1:AW309153	+	73.00	160.92	10.23	612	AW309153 sf94h02.y1 Gm-cl019 G
gb_gss:AZ233948	+	73.00	160.43	10.90	642	AZ233948 RPCI-23-73L24.TJ RPCI-
gb_gss:BH417892	+	73.00	157.97	14.94	815	BH417892 MusIt601-255 MuTAII-PC
gb_gss:BH373734	+	72.50	167.81	4.23	279	BH373734 AG-ND-126H18.TF ND-TAM
gb_gss:BH313283	+	72.50	164.07	6.83	401	BH313283 CH230-196N15.TJ CHORI-
gb_est2:BE591518	+	72.50	161.69	9.27	505	BE591518 kp47a07.y1 TBN95TM-SSR
gb_est1:AZ014621	+	72.50	159.60	12.11	618	AZ014621 RPCI-23-366D18.TJ RPCI
gb_gss:BH278768	+	72.50	159.29	12.60	637	BH278768 CH230-108N7.TV CHORI-2
gb_est2:BG524649	+	72.50	159.20	12.76	643	BG524649 5-44 Stevia field grow
gb_gss:AZ735738	+	72.50	157.59	15.67	751	AZ735738 RPCI-24-70J5.TJB RPCI-
gb_gss:AZ678077	+	72.50	156.63	17.74	835	AZ678077 ENTHK25TF Entamoeba h
gb_gss:AZ233624	+	72.50	156.50	18.03	825	AZ233624 ENPYC32TF Entamoeba h
gb_gss:AZ693407	+	72.50	155.57	20.32	914	AZ693407 ENIT04TF Entamoeba h
gb_gss:AZ441996	+	71.50	160.52	10.77	447	AZ441996 IM0234P21F Mouse 10kb
gb_gss:AZ233665	+	71.50	159.47	12.33	495	AZ233665 4908.gf21m18.sl Saccha
gb_est2:BE536132	+	71.50	155.94	19.38	687	BE536132 601061788F1 NIH_MGC_10
gb_gss:BH096764	+	71.50	155.17	21.40	751	BH096764 RPCI-24-287C31.TJ RPCI
gb_gss:CN505420	+	71.50	152.16	31.45	1005	AL320217 Tetraodon nigroviridi
gb_gss:AZ232971	+	71.00	158.46	14.02	485	AZ232971 IM0054F14F Mouse 10kb
gb_gss:AZ15059	+	71.00	158.46	14.02	485	AZ15059 IM0054F14F Mouse 10kb
gb_gss:BH600950	+	71.00	158.11	14.67	502	BH600950 B0G2V06TR B0G2 Brassic
gb_est2:BE442965	+	71.00	157.20	16.48	548	BE442965 WHE1108_G02_N04Z5 Whea
gb_est2:BM050293	+	71.00	153.75	25.66	766	BM050293 603632513F1 NIH_MGC_43
gb_gss:AZ672583	+	71.00	152.42	30.42	871	AZ672583 ENTJL17TR Entamoeba h

gb_est1:AW587961	+	70.50	158.30	14.32	438	AW587961 kp10c06.y1 TBN95TM-
gb_est1:AA823163	+	70.50	158.07	14.75	448	AA823163 v41e03.r1 SoaresM
gb_gss:BH295572	+	70.50	156.63	17.74	515	BH295572 CH230-187A17.TV CHO
gb_gss:AZ423995	+	70.50	156.24	18.65	535	AZ423995 IM0203B13F Mouse 10
gb_est2:BG804725	+	70.50	155.05	21.71	600	BG804725 0271-56 Mouse E14.5
gb_est2:BG808959	+	70.50	155.05	21.71	600	BG808959 2124-54 Mouse E14.5
gb_gss:BH061596	+	70.50	153.61	26.12	690	BH061596 RPCI-24-269B12.TJ R
gb_est2:BF295592	+	70.50	153.32	27.12	710	BF295592 025PB002 PB CDNA #1
gb_gss:BH665274	+	70.50	152.43	30.40	774	BH665274 602784057F1 NCI_CGA
gb_est2:BI695195	+	70.50	151.68	33.45	832	BI695195 603345846F1 NCI_CGA
gb_est2:BM454164	+	70.50	148.39	51.03	1145	BM454164 AGENCOURT_6419302
gb_est2:BM170235	+	70.50	153.32	27.11	631	BM170235 EST572758 PYBS Plas
gb_gss:AZ549155	+	70.00	151.89	32.58	725	AZ549155 RPCI-24-165I13.TV R
gb_est2:BI302215	+	69.50	155.87	19.55	438	BI302215 UI-R-DLO-clp-j-20.0
gb_gss:BH288978	+	69.50	153.35	26.99	559	BH288978 CH230-171N12.TVC CH
gb_est2:BE888235	+	69.50	150.42	39.32	743	BE888235 601511755F1 NIH_MGC
gb_gss:BH354216	+	69.50	150.05	41.23	770	BH354216 CH230-221H9.TJ CHOR
gb_gss:AZ549155	+	69.50	148.23	52.09	919	AZ549155 ENTMC48TR Entamoeba
gb_est1:AO69704	+	69.00	155.29	21.07	412	AO69704 AO069704 Rice panic
gb_est2:BF871591	+	69.00	154.04	24.72	465	BF871591 QVO-ET0148-O11100-4
gb_gss:BH354762	+	69.00	152.77	29.10	526	BH354762 CH230-199F18.TJ CHO
gb_gss:AZ944352	+	69.00	150.71	37.88	642	AZ944352 2M0205B19F Mouse 10
gb_gss:BH272068	+	69.00	148.69	49.08	781	BH272068 CH230-21K14.TV CHOR
gb_gss:BH132334	+	69.00	146.33	66.45	982	BH132334 ENTNY27TR Entamoeba
gb_est1:AI982549	+	68.50	153.85	25.33	421	AI982549 WT21H01.x1 NCI_CGAP
gb_est2:BM273517	+	68.50	150.41	39.40	588	BM273517 PFESToaas1908.y1 Pl
gb_gss:BH322134	+	68.50	149.01	47.11	673	BH322134 CH230-116C10.TV CHO
gb_gss:BH191484	+	68.50	149.00	47.20	674	BH191484 TC3-37A13.TR TC3 TI
gb_gss:BH352797	+	68.50	148.02	53.50	741	BH352797 CH230-33E6.TV CHORI
gb_gss:AZ698722	+	68.50	147.90	54.36	750	AZ698722 RPCI-23-232G24.TJ R
gb_gss:AZ671779	+	68.50	146.29	66.76	876	AZ671779 ENTJL09TF Entamoeba
gb_gss:AZ549109	+	68.50	146.04	68.98	898	AZ549109 ENTBJ05TR Entamoeba
gb_gss:AZ62869	+	68.50	145.73	71.74	925	AZ62869 ENTMS27TR Entamoeba
gb_gss:CN5071MR	+	68.50	145.07	78.06	986	CLONE BA0AB003D06 o
gb_est1:AO95147	+	68.00	153.01	28.21	406	AO95147 oy19a04.sl Soares s
gb_est2:BF014281	+	68.00	149.96	41.74	546	BF014281 PFESToaas75c03.y1 Pl
gb_est2:BM274078	+	68.00	149.66	43.37	562	BM274078 PFESToaas6e05.y1 Pl
gb_est1:AV835454	+	68.00	148.48	50.44	630	AV835454 AV835454 K. Sato un
gb_gss:AZ726577	+	68.00	148.40	50.97	635	AZ726577 RPCI-23-227G7.TJC R
gb_gss:AZ6808610	+	68.00	147.92	54.18	665	AZ6808610 602853163F1 NIH_MGC
gb_est2:BI088610	+	68.00	147.39	57.98	700	BI088610 602853163F1 NIH_MGC
gb_gss:BH296443	+	68.00	146.00	69.30	801	BH296443 CH230-123H12.TV CHO
gb_est2:BG023860	+	68.00	144.01	89.51	972	BG023860 602303507F1 NIH_MGC
gb_est1:AL514310	+	68.00	143.51	95.40	1020	AL514310 AL514310 LTI_NFL00
gb_est2:BM330986	+	67.50	150.05	41.25	481	BM330986 PIC1_68.B03.g1.A002
gb_est1:BE019434	+	67.50	149.92	41.93	487	BE019434 bb55e12.y1 NIH_MGC
gb_est1:AA786963	+	67.50	149.61	43.64	502	AA786963 m7308a1.fl Aspergill
gb_gss:BH310216	+	67.50	148.59	49.72	554	BH310216 CH230-3H2.TV CHORI-
gb_est1:BB625755	+	67.50	146.85	62.18	656	BB625755 BB625755 RIKEN full
gb_gss:BH191621	+	67.50	146.36	66.22	688	BH191621 TC3-37K13.TR TC3 TR
gb_gss:BH256660	+	67.50	145.44	74.49	755	BH256660 CH230-286G11.TJ CHO
gb_est2:BI856252	+	67.50	145.25	76.33	762	BI856252 603382985F1 NIH_MGC
gb_est2:BM041968	+	67.50	145.24	76.46	767	BM041968 603615846F1 NIH_MGC
gb_est2:BG120604	+	67.50	145.10	77.78	777	BG120604 6024364739F1 NIH_MGC
gb_est2:BE253581	+	67.50	144.38	85.28	833	BE253581 601109009F1 NIH_MGC
gb_est2:BG2585181	+	67.50	143.78	92.11	883	BG2585181 602358943F1 NCI_CGA
gb_est2:BG246099	+	67.50	143.73	92.80	888	BG246099 ENTSG95TF Entamoeba
gb_gss:BH159129	+	67.50	143.73	92.80	888	BH159129 ENTSG95TF Entamoeba
gb_gss:BH160705	+	67.50	143.67	93.50	893	BH160705 ENTSG95TF Entamoeba
gb_est2:BG288111	+	67.50	143.42	96.28	913	BG288111 602628811F1 NCI_CGA
gb_gss:AZ42954	+	67.50	142.95	102.46	957	AG420284 602484302F1 NIH_MGC
gb_est2:BM473346	+	67.50	142.68	102.46	957	BM473346 AGENCOURT_6466598
gb_hlc:BC004160	+	67.50	131.68	382.73	2592	BC004160 Homo sapiens, grow
gb_est2:D75662	+	67.00	151.82	32.85	360	D75662 CELK108DAF.Yu1 Kohar
gb_est1:BA458341	+	67.00	151.13	35.90	385	BA458341 sh86g05.y1 Gm-cl016
gb_est2:BE945386	+	67.00	149.75	42.84	440	BE945386 EST422965 MGHW Medi
gb_est2:BG730370	+	67.00	146.92	49.53	491	BG730370 df43a06.y1 Wellcome
gb_est2:BI448086	+	67.00	146.68	61.17	576	BI448086 dah91901.y1 NICHX
gb_gss:AZ007727	+	67.00	146.61	62.72	587	AZ007727 RPCI-23-246J3.TV RP
gb_est2:BI42651	+	67.00	146.61	64.14	597	BI42651 fq47h01.y1 Zebrafis
gb_est2:BI475809	+	67.00	146.55	64.56	600	BI475809 fq72h07.y3 Zebrafis
gb_est2:BI075556	+	67.00	146.55	64.56	600	BI075556 fsq4e09.y1 Zebrafis
gb_est2:BI841035	+	67.00	146.55	64.56	600	BI841035 fsq50d09.y1 Zebrafis



gb_est2:BI841074	67.00	146.55	64.56	600	! BI841074	fq50h06.y1 zebrafish a	gb_est1:AJ273964	572	1000	144.62	82.75	! AJ273964	AJ273964	Metarhiziu
gb_est2:BI841136	67.00	146.55	64.56	600	! BI841136	fq51f07.y1 zebrafish a	gb_gss:BH069552	587	383	66.00	144.35	! BH069552	RPCI-24-389M1.TVC R	
gb_est2:BI841165	67.00	146.55	64.56	600	! BI841165	fq52a06.y1 zebrafish a	gb_est1:BG815729	591	385	66.00	144.28	! BG815729	BD72e01.y1 Wellcom	
gb_est2:BI841250	67.00	146.55	64.56	600	! BI841250	fq53b04.y1 zebrafish a	gb_est1:AL649913	598	389	66.00	144.16	! AL649913	AL649913 XGC-gastru	
gb_est2:BI842618	67.00	146.55	64.56	600	! BI842618	fq47d07.y1 zebrafish a	gb_est2:BE942684	596	434	66.00	144.16	! BE942684	ECST-222623 MGHG Medi	
gb_est2:BI844454	67.00	146.55	64.56	600	! BI844454	fq34c07.y1 zebrafish a	gb_gss:AZ227993	626	503	66.00	143.69	! AZ227993	RPCI-23-79F21.TV RP	
gb_est2:BI865153	67.00	146.55	64.56	600	! BI865153	fq98a09.y1 zebrafish a	gb_gss:AL631463	646	507	66.00	143.69	! AL631463	EST562039 PVBs Plasm	
gb_est2:BI844464	67.00	146.54	64.71	601	! BI844464	fs34d05.y1 zebrafish a	gb_est1:AZ568742	660	514	66.00	143.14	! AZ568742	251Pvc10.y1 MBN #30	
gb_est2:CB4717	67.00	146.43	65.56	607	! CB4717	CB4717 dicystostellum dis	gb_gss:AO250262	672	700	66.00	142.96	! AO250262	TK17-Sp6 TAMU Arabi	
gb_est2:BI475552	67.00	146.40	65.85	609	! BI475552	fq67d08.y3 zebrafish a	gb_est1:AL628084	676	704	66.00	142.90	! AL628084	AL628084 XGC-gastru	
gb_est2:BI705474	67.00	146.40	65.85	609	! BI705474	fq73g11.y3 zebrafish a	gb_est1:AV406154	695	707	66.00	142.61	! AV406154	AV406154 Bombyx mor	
gb_est2:BI841170	67.00	146.40	65.85	609	! BI841170	fq52a11.y1 zebrafish a	gb_gss:BH319624	739	737	66.00	141.98	! BH319624	CH230-119P18.TVB CH	
gb_est2:BI845949	67.00	146.40	65.85	609	! BI845949	fq75c04.y1 zebrafish a	gb_est2:BG868866	752	752	66.00	141.80	! BG868866	60278434F1 NCI_CGA	
gb_est2:BI841264	67.00	146.35	66.85	612	! BI841264	fq53d12.y1 zebrafish a	gb_est2:BM159516	756	756	66.00	141.74	! BM159516	EST562039 PVBs Plasm	
gb_gss:AO928098	67.00	146.28	66.85	616	! AO928098	RPCI-23-258H6.TV RPCI	gb_gss:CMS07CGJ	761	761	66.00	141.68	! CMS07CGJ	T7 end of clone BD0	
gb_est2:BI475579	67.00	146.27	67.00	617	! BI475579	fq67g03.y3 zebrafish a	gb_gss:AO855490	774	774	66.00	141.50	! AO855490	CG61764B CPTOWAGDNA	
gb_est2:BI475662	67.00	146.27	67.00	617	! BI475662	fq68g12.y3 zebrafish a	gb_est1:AU133122	803	803	66.00	141.12	! AU133122	AU133122 NT2RP4 Hom	
gb_est2:BI475702	67.00	146.27	67.00	617	! BI475702	fq69d07.y3 zebrafish a	gb_est2:BF256770	936	936	66.00	139.54	! BF256770	HVSMER101P12f Hord	
gb_est2:BI845909	67.00	146.27	67.00	617	! BI845909	fq74f12.y1 zebrafish a	gb_est2:BG230505	978	978	66.00	139.09	! BG230505	602388366F1 NIH_MGC	
gb_gss:AZ447600	67.00	146.22	67.43	620	! AZ447600	IM0244P14R Mouse 10kb	gb_gss:BH266991	338	338	65.50	148.83	! BH266991	CH230-66J6.TV CHORI	
gb_est1:BE060308	67.00	146.17	67.86	623	! BE060308	HVSMER0011019f Hordeum	gb_est1:AU177894	357	357	65.50	148.27	! AU177894	AU177894 UV Irradia	
gb_est2:BI843687	67.00	146.12	68.29	626	! BI843687	fq48c12.y1 zebrafish a	gb_est1:AW636061	365	365	65.50	148.04	! AW636061	BI42c06.y1 lrracksh	
gb_est2:BG160282	67.00	145.99	69.45	634	! BG160282	fq53c08.y1 Wellcome CR	gb_gss:BH305635	493	493	65.50	148.94	! BH305635	CH230-37121.TV CHOR	
gb_gss:AZ417456	67.00	145.97	69.59	634	! AZ417456	IM0193F11R Mouse 10kb	gb_est1:AU036816	499	499	65.50	144.81	! AU036816	AU036816 Cryptomeri	
gb_est1:BI8412386	67.00	145.84	70.75	643	! BI8412386	AU123386 NT2RM2 Homo	gb_gss:AZ936889	505	505	65.50	144.69	! AZ936889	2M0193B23R Mouse 10	
gb_est2:BJ065577	67.00	145.54	73.53	662	! BJ065577	BJ065577 NIBB Mochii	gb_est2:BG230505	506	506	65.50	144.67	! BG230505	3prime NIH	
gb_gss:BH176803	67.00	145.39	75.01	672	! BH176803	007_H-20-rev SmaC1	gb_est1:AI469952	515	515	65.50	144.49	! AI469952	tj89a10.y1 Soares_N	
gb_gss:CMS07JHS	67.00	145.39	75.01	672	! CMS07JHS	T3_end of clone 007DD1	gb_gss:BH266198	528	528	65.50	144.23	! BH266198	CH230-55M10.TV CHOR	
gb_est2:BM183613	67.00	145.36	75.30	674	! BM183613	fv65e09.y1 Sugano SUD	gb_est2:BE678286	559	559	65.50	144.23	! BE678286	df53c06.y1 Xenopus	
gb_est1:AM18120	67.00	144.86	80.22	707	! AM18120	f108a09.y1 Sugano Kawa	gb_gss:BH029809	563	563	65.50	143.64	! BH029809	RPCI-24-324N23.TV R	
gb_gss:AZ06893	67.00	144.39	85.20	740	! AZ06893	RPCI-23-24612.TV RPCI	gb_est2:BI815335	591	591	65.50	143.57	! BI815335	PFEStoaa17H10.y1 P1	
gb_est1:AU091198	67.00	143.57	94.62	801	! AU091198	AU091198 lambda ZAPII	gb_est1:AW245971	604	604	65.50	142.84	! AW245971	2822766.3prime NIH	
gb_est1:AU005267	67.00	143.46	96.02	810	! AU005267	AU005267 Bombyx mori	gb_est2:BG016950	614	614	65.50	142.84	! BG016950	df68c11.y1 Xenopus	
gb_gss:AZ536055	67.00	143.35	97.44	819	! AZ536055	ENTIC022TR Entamoeba h	gb_gss:BH275351	624	624	65.50	142.67	! BH275351	CH230-35E3.TV CHORI	
gb_gss:AG672258	67.00	142.88	103.46	857	! AG672258	ENTIC022TR Entamoeba h	gb_gss:BH347225	624	624	65.50	142.51	! BH347225	CH230-42D16.TV CHOR	
gb_gss:BH148797	67.00	142.70	105.86	872	! BH148797	ENTPE74TF Entamoeba h	gb_gss:AO997223	640	640	65.50	142.25	! AO997223	RPCI-23-278L14.TV R	
gb_gss:AZ540419	67.00	142.48	108.93	891	! AZ540419	ENTCS55TR Entamoeba h	gb_gss:AZ734722	643	643	65.50	142.20	! AZ734722	RPCI-24-131P10.TV R	
gb_gss:BH148098	67.00	142.43	109.57	895	! BH148098	ENTPD95TR Entamoeba h	gb_gss:AZ245919	647	647	65.50	142.13	! AZ245919	2822950.3prime NIH	
gb_gss:BH136250	67.00	142.41	109.90	897	! BH136250	ENTNH55TR Entamoeba h	gb_gss:AZ636764	670	670	65.50	141.73	! AZ636764	1M0495116R Mouse 10	
gb_gss:BH155062	67.00	141.94	116.75	939	! BH155062	ENTSP09TR Entamoeba h	gb_gss:AO950992	700	700	65.50	141.32	! AO950992	Sheared DNA-52G22.T	
gb_gss:AG672290	67.00	141.89	117.41	943	! AG672290	ENTII07TR Entamoeba h	gb_gss:BH15764	704	704	65.50	141.26	! BH15764	RPCI-24-27114.TV RP	
gb_est1:BM475984	67.00	141.16	128.91	1012	! BM475984	AGENC00TR_6411302 NIH	gb_gss:BE792834	737	737	65.50	140.79	! BE792834	601584835F1 NIH_MGC	
gb_est2:CO6692	66.50	150.00	41.52	382	! CO6692	CO6692 Rat pancreatic is	gb_gss:AZ543858	901	901	65.50	138.72	! AZ543858	ENTH406TF Entamoeba	
gb_gss:AH267401	66.50	148.73	48.86	432	! AH267401	QV2-LR0051-130400-134	gb_gss:AZ689972	906	906	65.50	138.66	! AZ689972	ENTH406TF Entamoeba	
gb_gss:AZ2649188	66.50	148.49	51.37	442	! AZ2649188	RPCI-23-43C15.TJ RPCI	gb_gss:CMS079H3	906	906	65.50	138.66	! CMS079H3	T7 end of clone BV0	
gb_gss:AZ523113	66.50	148.31	51.56	450	! AZ523113	215PbE01 Pb MBN #21 P1	gb_gss:CMS06Y12	996	996	65.50	137.69	! CMS06Y12	T3 end of clone AB0	
gb_gss:AO442332	66.50	147.45	57.56	489	! AO442332	HS_5117-B1_C12-77A RPC	gb_gss:CMS071V6	1000	1000	65.50	137.64	! CMS071V6	clone BA0A8005G03	
gb_gss:AZ444932	66.50	147.39	58.03	492	! AZ444932	IM0240624R Mouse 10kb	gb_est1:AI987127	383	383	65.00	146.33	! AI987127	r522a05.y1 Sommer P	
gb_gss:BH306907	66.50	146.60	64.19	531	! BH306907	CH230-37M21.TJ CHORI-2	gb_est1:AW290550	385	385	65.00	146.27	! AW290550	NNXN031B06F NSF Xyl	
gb_gss:BH267401	66.50	145.91	70.17	568	! BH267401	CH230-82K21.TJ CHORI-2	gb_gss:AZ272800	389	389	65.00	146.17	! AZ272800	RPCI-23-161C17.TJ R	
gb_gss:AZ264780	66.50	145.89	70.33	569	! AZ264780	RPCI-23-156M11.TV RPCI	gb_gss:AO902030	434	434	65.00	145.04	! AO902030	HS_3009.A1_H09.T7 C	
gb_gss:BH332692	66.50	145.80	71.15	574	! BH332692	CH230-201O13.TV CHORI	gb_est2:BF198880	443	443	65.00	144.83	! BF198880	248865 MARC 2P1G Su	
gb_est2:BI21214025	66.50	145.69	72.13	580	! BI21214025	RE19713.Sprime RE Dros	gb_est1:AA227827	461	461	65.00	144.42	! AA227827	df56f12.r1 Soares_N	
gb_gss:BH278694	66.50	145.36	75.28	589	! BH278694	CH230-17119.TV CHORI-2	gb_est2:772291	465	465	65.00	144.33	! 772291	yc68e08.r1 Stratogens	
gb_gss:BH060138	66.50	145.17	77.28	611	! BH060138	RPCI-24-268L12.TV RPCI	gb_est1:AW290550	483	483	65.00	143.94	! AW290550	UO-HF-BMO-awh-h-11	
gb_gss:AZ5275565	66.50	144.37	85.41	659	! AZ5275565	307PvD12 Pb MBN #30 P1	gb_gss:BH322157	503	503	65.00	143.52	! BH322157	CH230-15A12.TV CHOR	
gb_est1:AA440960	66.50	144.26	86.61	666	! AA440960	LD12777.Sprime LD Dros	gb_est2:BF4651585	507	507	65.00	143.44	! BF4651585	U1-M-CG0P-bmu-c-07	
gb_est1:AA260343	66.50	144.03	89.20	681	! AA260343	LP04144.Sprime LP Dros	gb_est2:BI774580	514	514	65.00	143.29	! BI774580	466755 MARC 1B0V BO	
gb_gss:BH318347	66.50	143.90	90.76	690	! BH318347	CH230-119M20.TV CHORI	gb_est2:BE614138	536	536	65.00	142.86	! BE614138	601503882F1 NIH_MGC	
gb_gss:AZ859451	66.50	143.74	92.68	701	! AZ859451	2M0165M08F Mouse 10kb	gb_est2:BI791429	547	547	65.00	142.65	! BI791429	au05004.y1 Melton N	
gb_gss:BH335860	66.50	143.39	96.90	725	! BH335860	CH230-99H5.TV CHORI-23	gb_est1:AI929659	554	554	65.00	142.52	! AI929659	au03b05.y1 Schneide	
gb_est1:BB529740	66.00	155.00	21.85	209	! BB529740	BE597490 RIKEN full-le	gb_est2:BG041546	558	558	65.00	142.45	! BG041546	sv35g06.y1 Gm-cl057	
gb_est1:AV529706	66.00	150.74	37.75	316	! AV529706	AV529706 Arabidopsis t	gb_est2:BE251981	572	572	65.00	142.19	! BE251981	601107670F1 NIH_MGC	
gb_est2:BE507537	66.00	147.51	57.09	432	! BE507537	P11.70.G09.B1.A002 Pat	gb_est2:BE940929	593	593	65.00	141.82	! BE940929	EST420508 MGHG Medi	
gb_est2:BG507852	66.00	147.42	57.79	436	! BG507852	sac82N09.y1 Gm-cl072 C	gb_est1:AL046237	610	610	65.00	141.53	! AL046237	DKFZP434E057.r1 434	
gb_est1:AL902872	66.00	147.14	59.90	448	! AL902872	QV-BT019-200199-039 BT	gb_gss:AA550055	613	613	65.00	141.48	! AA550055	1145m3 gmbpFHB3.1	
gb_est2:BF648116	66.00	146.47	65.26	478	! BF648116	NF044E02ECIF1018 Elic	gb_est2:BG0505196	618	618	65.00	141.29	! BG0505196	602551663F1 NIH_MGC	
gb_est1:AL387541	66.00	145.80	71.10	510	! AL387541	MTBC43C07F1 MtBC Medic	gb_est2:BG041613	628	628	65.00	141.33	! BG041613	sv35g06.y1 Gm-cl057	
gb_est2:BM166865	66.00	145.43	74.63	529	! BM166865	EST569377 PVBs Plasmoc	gb_est2:BE886747	648	648	65.00	140.90	! BE886747	601506910F1 NIH_MGC	
gb_est2:BM168245	66.00	145.43	74.63	529	! BM168245	EST570768 PVBs Plasmoc	gb_est2:BE2887507	655	655	65.00	140.79	! BE2887507	601098576F1 NCI_CGA	
gb_est2:BG551921	66.00	144.99	78.95	552	! BG551921	dq22f06.y1 Wellcome CH	gb_est2:BG563573	659						



gb_est2:BF103873	65.00	140.48	140.65	675	1	BF103873	601647380F1 NIH_MGC_61	gb_gss:BH151875	64.50	136.41	237.27	891	1	BH151875	ENTPV13TR Entamoeba
gb_est2:BG499732	65.00	140.45	141.20	677	-	BG499732	602545149F1 NIH_MGC_60	gb_est2:BG499732	64.50	134.92	249.06	1029	1	BH151735	AGNCSQRT 6394897
gb_est2:BG298526	65.00	140.40	141.29	686	-	BG298526	CH230-84G12.TV CHORI-2	gb_gss:BM2500	64.50	133.40	349.07	1193	1	BM2500	F18D4-Sp6 IGF Arabid
gb_est1:AW689230	65.00	140.26	144.80	690	+	AW689230	NF016H08S9F1F1000 Devel	gb_est1:AV077298	64.00	147.78	55.20	263	1	AV077298	AV077298 Mus muscul
gb_est1:BF763780	65.00	140.26	144.80	690	+	BF763780	603049723F1 NIH_MGC_11	gb_est2:BF594749	64.00	147.54	83.62	360	1	BF594749	AS_c92_43E07_SKPLA
gb_est1:BB400791	65.00	139.86	152.34	717	-	BB400791	BB400791 RIKEN full-1e	gb_est2:BG062286	64.00	143.56	94.85	396	1	BG062286	LO550C11-5 NIA Mous
gb_est2:BF115568	65.00	139.58	157.99	737	-	BF115568	RPCI-24-386E1.TV RPCI	gb_est1:AW207979	64.00	143.17	99.64	411	1	AW207979	M11000e DSIR Medic
gb_gss:BH113568	65.00	139.29	163.97	758	-	BH132091	602981426F1 NCI_CGAP_1	gb_est2:BF126288	64.00	142.85	103.82	424	1	BF126288	NF034D06EC1F1058 El
gb_est2:BF191358	65.00	139.23	165.12	762	-	BF191358	603179137F1 NIH_MGC_12	gb_est2:BE943248	64.00	141.79	118.98	470	1	BE943248	EST422821 MGHG Medi
gb_est2:BF191195	65.00	139.11	167.70	771	-	BF791195	601584063F1 NIH_MGC_7	gb_est1:AL138698	64.00	141.74	119.65	472	1	AL138698	MEBC93C01F1 MLBC Me
gb_est1:AV762790	65.00	138.93	171.74	785	+	AV762790	AV762790 MDS Homo sapi	gb_est2:BI073353	64.00	141.72	119.98	473	1	BI073353	kt23h05.y1 Strongyl
gb_est1:AL120992	65.00	138.84	173.77	792	+	AU120992	AU120992 HEMBI Homo sa	gb_gss:PT017003U	64.00	141.49	123.69	484	1	PT017003U	parametum tetraure
gb_est2:BG330852	65.00	138.71	176.68	802	+	BG330852	60243149F1 NIH_MGC_18	gb_est2:BF1749607	64.00	141.05	130.84	505	1	BF1749607	ro80b09.y1 Heterode
gb_est2:BF1910892	65.00	138.59	179.30	811	-	BF1910892	60306766F1 NIH_MGC_11	gb_gss:AQ449743	64.00	140.91	133.24	512	1	AQ449743	500004Q0A.x1 CpiOwa
gb_gss:AX679937	65.00	138.30	186.06	834	-	AX679937	ENTMG61TR Entamoeba hi	gb_est2:BM285475	64.00	140.69	137.04	523	1	BM285475	524108 MARC 3BOV Ha
gb_est2:BF794797	65.00	138.14	189.90	847	-	BF794797	602256640F1 NIH_MGC_85	gb_est2:BF649673	64.00	140.49	140.52	533	1	BF649673	NF082B08EC1F1063 El
gb_est1:AL156518	65.00	137.99	193.77	860	-	AL156518	ALS56518 LTI_NFL006.PI	gb_est2:BF631721	64.00	140.47	140.86	534	1	BF631721	NF007A10DEC1F1070 Dr
gb_gss:AX686076	65.00	137.96	194.37	862	-	AX686076	ENTIN73TR Entamoeba hi	gb_est2:BF647076	64.00	140.45	141.21	535	1	BF647076	NF005A10DEC1F1070 El
gb_est2:BF668207	65.00	137.56	204.57	896	-	BF668207	ENTLV61TR Entamoeba hi	gb_gss:BH049123	64.00	140.26	144.71	545	1	BH049123	RPCI-24-39G120.TV R
gb_gss:AX668951	65.00	137.55	204.87	897	-	AX668951	ENTFW25TR Entamoeba hi	gb_est2:BG448038	64.00	139.76	154.27	572	1	BG448038	NF105B02EC1F1015 El
gb_gss:AX671401	65.00	137.36	210.02	914	-	AX671401	ENTIB51TR Entamoeba hi	gb_est2:BG943248	64.00	139.66	155.70	576	1	BG943248	NF095260 N1BB Mochi
gb_gss:BH151832	65.00	137.34	210.63	916	-	BH151832	ENTPS50TR Entamoeba hi	gb_est2:BG512963	64.00	139.66	156.42	578	1	BG512963	da64h09.y1 Wellcom
gb_est2:BG823010	65.00	137.26	212.76	923	-	BG823010	602728030F1 NIH_MGC_15	gb_est2:BG447794	64.00	139.62	157.13	580	1	BG447794	NF070D10EC1F1079 El
gb_gss:AX686785	65.00	137.22	213.68	926	-	AX686785	ENTIC71TR Entamoeba hi	gb_est2:BG452134	64.00	139.48	160.01	588	1	BG452134	NF077A04FL1F1033 De
gb_est2:BF1912052	65.00	137.16	215.51	932	-	BF1912052	601663229F1 NCI_CGAP_1	gb_est2:BE5941470	64.00	139.34	162.89	596	1	BE5941470	EST421049 MGHG Medi
gb_gss:CN506WRS	65.00	137.09	217.35	938	-	BN886519	601508248F1 NIH_MGC_71	gb_gss:BM443219	64.00	139.32	163.25	597	1	BM443219	BOGNI25TR BOGN Bras
gb_est2:BF139227	65.00	136.83	224.73	962	-	AL148814	T3 end of clone AX0AAC	gb_est2:BE941610	64.00	139.05	169.06	613	1	BE941610	EST421189 MGHG Medi
gb_est2:BM456645	65.00	136.29	240.94	1014	+	BF139227	601783383F1 NCI_CGAP	gb_est2:BF641945	64.00	139.03	169.43	614	1	BF641945	EST421461 MGHG Medi
gb_est2:BE876066	65.00	135.71	259.33	1070	+	BM456645	AGNCSQRT_6408876 NIH	gb_est2:BF649446	64.00	139.00	170.16	616	1	BF649446	NF04487134F1 NIH_MGC
gb_est2:BM462173	65.00	135.62	262.54	1082	-	BE876066	601485749F1 NIH_MGC_6	gb_est2:BF649328	64.00	138.95	171.26	619	1	BF649328	NF0875686 F01487134F1 NIH_MGC
gb_est2:BM465712	65.00	135.35	271.56	1110	-	BM462173	AGNCSQRT_6424373 NIH	gb_est2:BF646143	64.00	138.51	181.21	646	1	BF646143	NF0875686 F01487134F1 NIH_MGC
gb_est2:BM458555	65.00	135.15	278.70	1132	-	BM465712	AGNCSQRT_6484795 NIH	gb_est2:BF649328	64.00	138.41	183.44	652	1	BF649328	NF0875686 F01487134F1 NIH_MGC
gb_est2:BG339650	65.00	135.12	279.68	1135	-	BM458555	AGNCSQRT_6414130 NIH	gb_est2:BG543541	64.00	138.33	185.30	657	1	BG543541	NF097F08FL1F1064 De
gb_est2:BF489462	65.00	134.95	285.89	1154	-	BG339650	602437213F1 NIH_MGC_4	gb_est2:BG543541	64.00	138.33	185.30	657	1	BG543541	NF097F08FL1F1064 De
gb_htc:BC017057	65.00	130.93	479.09	1705	+	BF489462	603021023F1 NIH_MGC_1	gb_est2:BG543541	64.00	138.33	185.30	657	1	BG543541	NF097F08FL1F1064 De
gb_est2:BG508200	64.50	146.04	68.95	350	-	BC017057	Homo sapiens, Similar	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est1:BB761684	64.50	143.16	99.82	463	-	BM508200	Nb_ad1.196.M1 Nippo	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BG500175	64.50	143.09	100.68	466	-	BM508200	sac99g03.y1 Gm-cl073 C	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX0500175	64.50	142.36	110.51	500	+	BF761684	BB761684 RIKEN full-1e	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX295038	64.50	142.02	115.50	517	-	AX0500175	V41D6 mtN-3xHA/lacZ GI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX912250	64.50	141.63	121.45	537	-	AX295038	CH230-104A17.TV CHORI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BM350930	64.50	141.55	122.65	541	-	AX912250	RPCI-24-167D24.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM106317	64.50	141.44	124.45	547	-	BM350930	MESP270-G11.T3 ISUMS-F	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM065271	64.50	141.20	128.38	560	-	BM106317	RPCI-24-306A13.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX394444	64.50	141.05	130.81	568	-	BM065271	RPCI-24-285H24.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX796671	64.50	140.71	132.94	575	-	AX394444	LM0158G04F Mouse 10kb	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX796671	64.50	140.68	137.24	589	-	AX796671	LM0052804R Mouse 10kb	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM341635	64.50	140.64	137.86	591	-	AX796671	RPCI-23-225D23.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX158181	64.50	140.62	138.17	592	-	BM341635	CH230-183G21.TV CHORI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX2824153	64.50	140.31	143.75	612	-	AX158181	nbxd001L0D18F CUGI Rice	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX094374	64.50	140.26	144.69	613	-	AX2824153	CH230-166E11.TV CHORI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX2601985	64.50	140.08	148.13	624	-	AX094374	RPCI-23-315O10.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX2004955	64.50	140.06	148.44	625	-	AX2601985	LM0420C06R Mouse 10kb	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BG6292327	64.50	140.05	148.76	626	-	AX004955	RPCI-23-375A1.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BF645902	64.50	140.05	148.76	626	-	BG6292327	HNC601-C11.R HNC (Hum	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM037806	64.50	139.90	151.59	635	-	BF645902	NF042D04EC1F1041 Ellic	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM067137	64.50	139.80	153.49	641	-	BM037806	RPCI-24-387L7V RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM092858	64.50	139.72	155.08	646	-	BM067137	RPCI-24-371C9.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est1:BM574242	64.50	139.69	155.71	648	-	BM092858	RPCI-24-375C10.TV RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BF027885	64.50	139.36	162.42	669	-	BM574242	ESR316833 GVN Medicag	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_est2:BF006345	64.50	139.32	163.39	672	-	BF027885	BB127885 RIKEN full-1e	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:AX2040313	64.50	138.43	182.95	732	-	BF006345	EST434843 DSLC Medicag	gb_gss:AX2040313	64.50	138.30	186.12	586	1	AX2040313	RPCI-23-233N23.TV R
gb_gss:BM347912	64.50	138.36	184.61	737	-	AX2040313	CH230-243Z02.TV CHORI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM151895	64.50	137.74	200.00	783	-	BM347912	RPCI-23-243E17V RPCI	gb_est2:BF645768	64.00	137.80	196.57	692	1	BF645768	NF017C04EC1F1033 El
gb_gss:BM263445	64.50	137.62	203.05	792	-	BM151895	BOHEB30TF BOHE Brassic	gb_est2:BF645768	64.00						



gb_est1:AW921289	63.50	137.20	214.34	652	! AW921289	EST352593	Rat gene ind	gb_gss:AO766307	62.50	139.53	158.96	411	! AO766307	HS_5489_B2_D06_T7A
gb_est1:AA056151	63.50	137.07	217.83	660	! AA056151	zfs5e09.1	Pv MBN #30	gb_est2:BC405800	62.50	139.50	159.47	412	! BC405800	SC38h11.y1 Gm-cl05
gb_gss:AZ573008	63.50	137.07	217.83	660	gb_est1:312pve11	rv Soares rct1		gb_est2:BC725933	62.50	139.36	162.55	418	! BC725933	BB25933 RIKEN full
gb_est1:BB175228	63.50	137.00	220.01	665	! BB175228	BB175228	RIKEN full-1e	gb_est2:BG653982	62.50	139.36	162.55	418	! BG653982	sa063q01.y2 Gm-cl05
gb_gss:AZ095092	63.50	136.86	223.96	674	! AZ095092	RPCI-23-461D5	.TV RPCI-1	gb_est2:BG663435	62.50	139.04	169.27	431	! BG663435	DRAAQA12.Rat DRG-L
gb_gss:AZ2824897	63.50	136.84	224.40	675	! AZ2824897	2M0099117R	Mouse 10kb	gb_est1:AG5053378	62.50	138.53	180.79	453	! AG5053378	UI-HF-BN0-alu-e-06-
gb_est2:BF165454	63.50	136.75	227.04	681	! BF165454	60177345F1	NCI_CGAP-1	gb_est1:BE059860	62.50	138.35	185.03	461	! BE059860	sn38b02.y1 Gm-cl01e
gb_est2:BF137660	63.50	136.73	227.48	682	! BF137660	601783189F1	NCI_CGAP-1	gb_est2:HG303030	62.50	138.19	188.75	468	! HG303030	yy06B04.s1 Soares fet
gb_gss:AZ236356	63.50	136.69	228.81	685	! AZ236356	1M0103N04R	Mouse 10kb	gb_gss:BH3134333	62.50	137.93	195.18	480	! BH3134333	CH230-101M12.TJ CHO
gb_gss:AZ3135427	63.50	136.64	230.13	688	! AZ3135427	BB135427	RIKEN full-1e	gb_gss:AZ321786	62.50	137.93	195.18	480	! AZ321786	1M00420H08R Mouse 10
gb_gss:AZ2876142	63.50	136.64	230.13	688	! AZ2876142	2M0191H06F	Mouse 10kb	gb_gss:BH296566	62.50	137.93	195.18	480	! BH296566	CH230-92N14.TJB CHO
gb_gss:AZ2936570	63.50	136.61	231.02	690	! AZ2936570	2M0193N13F	Mouse 10kb	gb_gss:AZ2909362	62.50	137.70	201.12	491	! AZ2909362	1M0434002F Mouse 10
gb_gss:AZ2958916	63.50	136.54	231.24	695	! AZ2958916	2M0226G22F	Mouse 10kb	gb_gss:BH366225	62.50	137.70	201.12	491	! BH366225	CH230-28K9.TJ CHORI
gb_gss:BH381735	63.50	136.42	236.79	703	! BH381735	AG-ND-16466	.TR ND-TAM	gb_gss:AO327844	62.50	137.55	203.83	496	! AO327844	RPCI-23-262A4.TV RP
gb_gss:AZ487153	63.50	136.38	238.13	706	! BG585389	EST487153	MHAM Medica	gb_gss:AG663209	62.50	137.55	204.92	498	! AG663209	1M0542A18R Mouse 10
gb_gss:BH311576	63.50	136.15	245.29	722	! BH311576	CH230-62K6	.TV CHORI-23	gb_gss:AZ076709	62.50	137.32	210.93	509	! AZ076709	RPCI-23-409P8.TV RP
gb_gss:AZ225664	63.50	136.15	245.29	722	! BH322664	CH230-7G11	.TVB CHORI-2	gb_gss:BH258587	62.50	137.22	210.93	514	! BH258587	CH230-75K9.TJ CHORI
gb_gss:AZ282673	63.50	136.05	248.44	729	! BH328273	AG-ND-143J18	.TF ND-TAM	gb_est1:AA942620	62.50	137.14	215.87	518	! AA942620	LDX6942.Sprime LD D
gb_gss:AZ1660780	63.50	135.34	272.15	781	! BV660780	603303821F1	NIH_CGAP-M	gb_gss:AZ451320	62.50	137.08	217.53	521	! AZ451320	1M0250H08R Mouse 10
gb_est1:AV384514	63.50	135.30	273.53	784	! AV384514	AV384514	Halocynthia x	gb_est2:BI362580	62.50	136.77	226.41	537	! BI362580	RE47178.Sprime RE D
gb_gss:BH354354	63.50	134.52	301.09	843	! BH354354	CH230-221K12	.TJ CHORI-1	gb_est1:AW851247	62.50	136.53	233.69	550	! AW851247	IL3-CY0220-160200-0
gb_gss:BH437532	63.50	134.52	302.03	845	! BH437532	BOGJ35TR	BOGI Brassic	gb_est1:AW851403	62.50	136.53	233.69	550	! AW851403	IL3-CY0220-170200-0
gb_gss:AZ2683120	63.50	134.19	315.34	873	! AZ2683120	ENTIG94TF	Entamoeba hi	gb_gss:AZ591483	62.50	136.47	235.37	553	! AZ591483	1M0401A11R Mouse 10
gb_gss:CN504650	63.50	133.93	325.89	895	! AL290055	Tetraodon nigroviridis		gb_gss:AZ238114	62.50	136.17	244.99	569	! AZ238114	RPCI-23-84I23.TV RP
gb_gss:AZ5026887	63.50	133.75	333.62	911	! AL26887	AL526887	ITI_NFL003	gb_gss:AZ949436	62.50	136.16	244.99	570	! AZ949436	2M0212L19R Mouse 10
gb_est2:BG414813	63.50	133.73	334.59	925	! BE542931	601068068F1	NIH_MGC-10	gb_gss:AZ579605	62.50	136.12	246.13	572	! AZ579605	1M0367A02R Mouse 10
gb_est1:BG414813	63.50	133.59	340.42	925	! BG414813	HVSMK003N02F	Hordeum	gb_gss:AZ779972	62.50	135.98	250.69	580	! AZ779972	2M0016D21R Mouse 10
gb_est1:BB404122	63.00	143.49	95.69	315	! BB404122	1M040122	RIKEN full-1e	gb_est1:AU023276	62.50	135.92	252.41	583	! AU023276	Mouse unfe
gb_gss:AZ595361	63.00	143.29	98.11	321	! AZ595361	BM04071221R	Mouse 10kb	gb_est1:BB608819	62.50	135.91	252.98	584	! BB608819	RIKEN full
gb_est1:BB114379	63.00	142.91	102.99	333	! BB114379	BB114379	RIKEN full-1e	gb_gss:AZ512013	62.50	135.91	252.98	584	! AZ512013	1M0357M14F Mouse 10
gb_gss:BH364962	63.00	139.49	159.71	464	! BH364962	CH230-75C13	.TJ CHORI-2	gb_gss:AZ719908	62.50	135.75	258.15	593	! AZ719908	RPCI-24-104G16.TV R
gb_est1:AW639312	63.00	138.91	172.12	491	! AW639312	bl80c09.w1	Blackshari	gb_gss:AZ377703	62.50	135.73	258.15	593	! AZ377703	1M0132U12F Mouse 10
gb_gss:AA945633	63.00	138.89	172.58	492	! AA945633	C049 Zhou and Ragan	19	gb_gss:AZ655188	62.50	135.71	259.30	595	! AZ655188	1M0503020F Mouse 10
gb_gss:AZ292315	63.00	138.64	178.17	504	! AZ292315	2M0276K24R	Mouse 10kb	gb_gss:AZ231092	62.50	135.63	262.19	600	! AZ231092	RPCI-23-53D9.TV R
gb_gss:BH351415	63.00	138.36	184.75	518	! BH351415	CH230-175A7	.TV CHORI-2	gb_gss:AZ738995	62.50	135.61	262.77	601	! AZ738995	RPCI-24-102M12.TJ R
gb_gss:BH315580	63.00	137.95	194.72	539	! BH315580	CH230-194B19	.TV CHORI-2	gb_gss:AZ115406	62.50	135.59	263.34	602	! AZ115406	RPCI-23-90222.TV R
gb_gss:AZ2965858	63.00	137.81	198.07	546	! AZ2965858	2M4PVC09	Pv MBN #16 (a	gb_gss:BH083552	62.50	135.54	265.08	603	! BH083552	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_est2:BI575390	62.50	135.08	281.43	633	! BI575390	RH31011.Sprime RH D
gb_gss:AV900606	63.00	136.18	244.38	640	! AV900606	AV900606	Nori Satoh un	gb_gss:AZ949975	62.50	134.98	284.96	639	! AZ949975	2M0213N15R Mouse 10
gb_est1:BB630829	63.00	136.16	244.88	641	! BB630829	BB630829	RIKEN full-1e	gb_gss:BH107746	62.50	134.98	284.96	639	! BH107746	RPCI-24-352E14.TV R
gb_gss:AZ226665	63.00	136.13	245.89	643	! BE226665	ia2ae03.y1	Mouse E10 5	gb_gss:BH040260	62.50	134.96	285.55	640	! BH040260	RPCI-24-329J10.TV R
gb_gss:BH309516	63.00	135.89	253.51	658	! BH309516	CH230-38K11	.TJ CHORI-2	gb_gss:AZ002247	62.50	134.87	289.10	646	! AZ002247	RPCI-23-340L23.TV R
gb_gss:AZ295558	63.00	135.83	255.55	662	! Z91518.F	rubripes GSS sequence		gb_gss:AZ510507	62.50	134.82	290.87	649	! AZ510507	1M035SK06F Mouse 10
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50	135.54	265.08	603	! AZ562245	RPCI-24-293E22.TJ R
gb_gss:AZ2971127	63.00	136.40	242.36	636	! AZ2971127	2M0244F01R	Mouse 10kb	gb_gss:AZ562245	62.50					



gb_gss: A0411905	62.50	132.88	372.84	783	I A0411905 CPG0901A CpiOWAGDnAl C	gb_est2:BI600453	62.00	127.74	721.02	1146	I BI600453 603246492F1 NIH_MG
gb_gss: BH283394	62.50	132.87	373.47	784	I BH283394 CH230-206A23 TJ CHORI-	gb_gss: A2774533	61.50	142.74	105.37	276	I A2774533 2M00A0N0F Mouse 10
gb_est2:BF191613	62.50	132.78	377.89	791	I BF191613 602713990F1 NIH_MGC_48	gb_est1:BE100855	61.50	141.21	128.18	238	I BE100855 UI-R-BJ1-atx-a-11-0
gb_gss: A2675661	62.50	132.04	415.61	850	I A2675661 ENTHV05TR Entamoeba h	gb_est1:AW698531	61.50	140.95	132.49	283	I AW698531 g326 glandular-hair
gb_gss: BH164326	62.50	131.29	457.50	914	I BH164326 ENTHV071TR Entamoeba h	gb_est1:BE024048	61.50	140.52	139.97	235	I BE024048 sm96a08-y1 Gm-cl015
gb_gss: A2669831	62.50	130.78	488.20	960	I A2669831 ENTHV05TR Entamoeba h	gb_gss: A2325587	61.50	138.44	182.82	361	I A2325587 EPT568330 PyBS Plas
gb_gss: CNS0662G	62.50	130.46	508.47	990	I AL422546 T7 end of clone A20A0A	gb_est2:BM165807	61.50	137.72	200.43	387	I BM165807 EPT568330 PyBS Plas
gb_hcc: AK010578	62.50	130.24	523.47	1012	I AL397982 T3 end of clone A20A0A	gb_gss: A2803117	61.50	137.48	206.62	336	I A2803117 2M00G6318F Mouse 10
gb_hcc: AK014389	62.50	120.56	1.8e+03	2587	I AK010578 Mus musculus ES cells	gb_est2:R642223	61.50	137.46	207.31	397	I R642223 y118g03 r1 Soares pla
gb_est1:AV720589	62.50	117.42	2.7e+03	3506	I AK014389 Mus musculus adult ma	gb_est2:BM171135	61.50	137.30	211.47	403	I BM166383 EPT568906 PyBS plas
gb_gss: A0874459	62.00	141.32	126.28	307	I AV220589 AV220589 RIKEN full-le	gb_est2:BM171135	61.50	137.23	213.55	406	I BM171135 EPT573658 PyBS Plas
gb_gss: A0874459	62.00	138.24	187.54	414	I A0874459 LMAJFV1-Lm45B04-y1 Lei	gb_gss: A2758002	61.50	137.13	216.34	410	I A2758002 IM0551D02F Mouse 10
gb_est2:BF399383	62.00	137.95	194.76	426	I BF399383 WHE00335 E05F990516 ITR	gb_gss: A2760215	61.50	136.59	231.83	432	I A2760215 SP-0065-B1-E10-SP6E
gb_est2:BF191613	62.00	137.52	205.72	444	I BF514515 UI-H-BW1-anf-g-09-U01	gb_est1:AW720358	61.50	136.45	236.09	438	I AW720358 LjNESP2169r Lotus J
gb_gss: A2505436	62.00	137.09	217.44	463	I A2505436 LM0346B08F Mouse 10kb	gb_gss: BH287498	61.50	136.33	239.66	443	I BH287498 CH230-180P22 TJ CHO
gb_est1:AW051431	62.00	136.87	223.68	473	I AW957405 EPT369595 MAGE rescue	gb_gss: BH339590	61.50	136.19	243.97	449	I BH339590 CH230-130H5 TJ CHOR
gb_est1:AA308815	62.00	136.42	236.90	494	I AA308815 EPT197968 Colon carcino	gb_gss: BH332971	61.50	136.01	249.73	457	I BH332971 CH230-736A TJ CHORI
gb_gss: A0874459	62.00	136.34	239.44	498	I A0874459 LMAJFV1-Lm45B04-y1 Lei	gb_gss: A2925066	61.50	135.85	254.81	484	I A2925066 4910.ez31m02.s1 Sac
gb_est2:BF399383	62.00	136.31	240.08	499	I BF399383 WHE00335 E05F990516 ITR	gb_gss: A2490839	61.50	135.78	256.99	487	I A2490839 IM0324L02F Mouse 10
gb_gss: A2603986	62.00	136.21	243.27	504	I A2603986 LM0423B12R Mouse 10kb	gb_est1:AL635732	61.50	135.31	273.12	489	I AL635732 ALG35732 XGC-neurul
gb_gss: BH304357	62.00	136.09	247.11	510	I BH304357 CH230-200P4 TV CHORI-2	gb_gss: AQ629542	61.50	135.27	274.60	491	I AQ629542 RPCI-11-469M14 TJ R
gb_gss: AQ181449	62.00	135.91	252.89	519	I AQ181449 HS 3218-B1-F02-T7 CIT	gb_est1:AL117988	61.50	135.16	278.30	496	I AL117988 uc42d04 r1 Soares-m
gb_est2:BF399383	62.00	135.75	258.06	527	I BF399383 WHE00335 E05F990516 ITR	gb_gss: BH266294	61.50	135.14	279.04	497	I BH266294 CH230-558R TV CHORI
gb_est2:BF399383	62.00	135.39	270.43	546	I BF399383 WHE00335 E05F990516 ITR	gb_est2:BF080202	61.50	135.12	279.79	498	I BF080202 602876618F1 NCI-CGA
gb_est2:BF399383	62.00	135.37	271.09	547	I BF191613 602713990F1 NIH_MGC_48	gb_gss: A2992501	61.50	134.94	286.50	507	I A2992501 M02771L10F Mouse 10
gb_est2:BF399383	62.00	135.16	278.32	558	I BF176967 EM1-4_H09-B1-A002 Embri	gb_gss: AQ974442	61.50	134.90	287.99	509	I AQ974442 RPCI-23-315P3 TJ RP
gb_est2:BF399383	62.00	135.03	282.95	565	I BF399383 WHE00335 E05F990516 ITR	gb_est2:BF1436831	61.50	134.82	290.99	513	I BF1436831 gc53a06 y1 Moss EST
gb_gss: A2812706	62.00	135.00	284.28	567	I A2812706 2M0079M04R Mouse 10kb	gb_est2:BM160781	61.50	134.74	293.99	517	I BM160781 EPT563304 PyBS Plas
gb_gss: A2923157	62.00	134.73	294.27	582	I A2923157 RPCI-23-291N6 TJ RPCI-	gb_est2:BF201021	61.50	134.64	297.76	522	I BF201021 BJ201021 normalized
gb_est1:AW761303	62.00	134.59	299.63	590	I AW761303 s165g06-y1 Gm-cl027 GI	gb_gss: A2771471	61.50	134.64	297.76	522	I A2771471 IM0573H1R Mouse 10
gb_est2:BF191613	62.00	134.50	302.99	595	I BF191613 602713990F1 NIH_MGC_48	gb_est2:BF206472	61.50	134.60	299.27	524	I BF206472 BJ206472 normalized
gb_est2:BF191613	62.00	134.41	306.36	600	I BE847280 LjNESP55C11r Lotus jag	gb_est2:BF108080	61.50	134.58	300.02	525	I BF108080 pg11n.pk005.h7 Norm
gb_est2:BF191613	62.00	134.28	311.78	608	I A2107300 RPCI-23-456H16 TJ RPCI-	gb_est2:BF203608	61.50	134.58	300.02	525	I BF203608 BJ203608 normalized
gb_gss: A2107300	62.00	134.26	312.46	609	I BE225532 MD041-Meloidogyne inc	gb_est2:BF191738	61.50	134.56	300.78	526	I BF191738 BJ191738 normalized
gb_est1:BE225532	62.00	134.04	321.31	622	I BH308115 CH230-4K33 TV CHORI-23	gb_est2:BF207024	61.50	134.56	300.78	526	I BF207024 BJ207024 normalized
gb_gss: BH308115	62.00	134.01	322.68	624	I BF308115 s336g02-y1 Gm-cl028 GI	gb_est2:BF033727	61.50	134.40	306.85	534	I BF033727 EPT432225 KVI Medici
gb_est1:AW761303	62.00	133.75	333.66	640	I BF859863 963014G12.x2 C. reinha	gb_gss: A2377491	61.50	134.38	309.13	535	I A2377491 IM0131D22R Mouse 10
gb_est1:AW761303	62.00	133.73	334.35	641	I A237457 UI132457 NT2RP3 Homo s	gb_gss: CNS060CH2	61.50	134.34	309.13	537	I AL392541 T7 end of clone AR0
gb_gss: A275682	62.00	133.73	334.35	641	I B61244 T2114TR TAMU Arabidopsis	gb_est1:AL639271	61.50	134.31	310.65	539	I AL639271 BJ639271 normalized
gb_gss: A275682	62.00	133.62	339.19	648	I A275682 2M0250P16R Mouse 10kb	gb_est1:AL639271	61.50	134.19	315.23	545	I AL639271 AL629271 XGC-gastru
gb_gss: BH286229	62.00	133.45	346.83	659	I BH286229 CH230-206I19 TJ CHORI-	gb_est2:BF202045	61.50	133.91	326.76	560	I BF202045 BJ193246 normalized
gb_est2:BF191613	62.00	133.32	352.41	667	I B5143288 SIC-B08 Sugar Beet str	gb_est2:BF193246	61.50	133.91	326.76	560	I BF193246 BJ193246 normalized
gb_est1:AW761303	62.00	132.71	381.34	709	I AW761303 s165g06-y1 Gm-cl027 GI	gb_est2:BF195265	61.50	133.91	326.76	560	I BF195265 BJ195265 normalized
gb_est1:BF1436831	62.00	132.55	389.20	719	I BF1436831 Gm01.10b02 R Gm01-AAFC	gb_est2:BF195265	61.50	133.91	326.76	560	I BF195265 BJ195265 normalized
gb_est2:BF399383	62.00	132.51	391.34	722	I BG839231 Gm01.10b02 R Gm01-AAFC	gb_est2:BF197742	61.50	133.89	327.53	561	I BF197742 BJ197742 normalized
gb_est2:BF399383	62.00	132.46	393.50	725	I BJ149430 BJ149430 unpublshed c	gb_gss: AQ7092653	61.50	133.71	335.28	571	I AQ7092653 HS_5421_B2_H07-A7A
gb_gss: A2898504	62.00	132.46	393.50	725	I A2898504 RPCI-24-186624 TV RPCI-	gb_gss: A2350262	61.50	133.66	337.61	574	I A2350262 IM0087B17R Mouse 10
gb_est1:AW722384	62.00	132.36	398.53	732	I A722384 fc26f01.x1 Zebrafish w	gb_est2:BF196283	61.50	133.57	341.50	579	I BF196283 BJ196283 normalized
gb_est1:AW706679	62.00	132.02	416.63	757	I AW706679 AV706679 ADB Homo sap	gb_est2:BF1436484	61.50	133.53	343.07	581	I BF1436484 9C47b02.y1 Moss EST
gb_est1:AW706679	62.00	131.96	419.55	761	I AW706679 AV706679 ADB Homo sap	gb_gss: A2705220	61.50	133.39	349.33	589	I A2705220 RPCI-23-226K4 TV RP
gb_gss: A2705220	62.00	131.73	431.99	778	I AL425755 clone BAOAB007G06 of l	gb_est2:BF200119	61.50	133.27	354.83	596	I BF200119 BJ200119 normalized
gb_est1:AW706679	62.00	131.71	433.46	780	I A218401 B02H8401 unpublshed c	gb_est2:BF197459	61.50	133.25	355.62	597	I BF197459 BJ197459 normalized
gb_gss: A2705220	62.00	131.71	433.46	780	I B4511651 B02H8401 unpublshed c	gb_est2:BF191116	61.50	133.20	357.98	600	I BF191116 BJ191116 normalized
gb_gss: A2705220	62.00	131.52	443.78	794	I B426932 B0H438TF BOH4 Brassic	gb_est2:BF198208	61.50	133.20	357.98	600	I BF198208 BJ198208 normalized
gb_est1:AW706679	62.00	131.34	454.15	808	I A2125115 A2125115 NT2RM4 Homo s	gb_est2:BF196284	61.50	133.11	361.93	605	I BF196284 BJ196284 normalized
gb_est1:AW706679	62.00	131.05	471.33	831	I BF212744 601813987F1 NIH_MGC_54	gb_gss: BH283936	61.50	133.05	365.10	609	I BH283936 CH230-192B10 TJ CHO
gb_gss: A2705220	62.00	130.82	485.64	849	I A212640 A212640 Y79AA1 Homo s	gb_gss: BH313775	61.50	133.00	367.48	612	I BH313775 CH230-192B10 TJ CHO
gb_gss: A2705220	62.00	130.82	485.64	850	I BM359139 GA_Ea0016P23r Gossyp	gb_est1:AL967425	61.50	132.98	368.28	613	I AL967425 Ljlrnpest04-320-d9
gb_gss: A2705220	62.00	130.75	490.18	856	I A212640 A212640 Y79AA1 Homo s	gb_gss: A2904071	61.50	132.96	369.07	614	I A2904071 RPCI-24-161C8 TV RP
gb_est1:AW706679	62.00	130.70	493.21	860	I A212640 A212640 Y79AA1 Homo s	gb_gss: A2904071	61.50	132.96	369.07	614	I A2904071 RPCI-24-161C8 TV RP
gb_gss: A2705220	62.00	130.52	504.62	875	I A2544134 ENTHV068TR Entamoeba h	gb_est2:BF167811	61.50	132.89	372.25	618	I BF167811 EPT570334 PyBS Plas
gb_est1:AW706679	62.00	130.29	519.93	895	I AL544384 A2544384 LTI_NFH.MGC.54	gb_est1:AL640635	61.50	132.86	373.85	620	I AL640635 AL640635 XGC-neurul
gb_est2:BF191613	62.00	130.12	531.49	910	I BG613944 602641662F1 NIH_MGC.61	gb_gss: A2662278	61.50	132.83	375.44	622	I A2662278 IM0541H17F Mouse 10
gb_gss: A2705220	62.00	130.05	536.13	916	I AL417621 T3 end of clone AX0A0A	gb_gss: A2734345	61.50	132.42	395.53	647	I A2734345 LM0127D05F Mouse 10
gb_gss: A2705220	62.00	129.81	553.22	938	I BE536005 601062437F1 NIH_MGC.10	gb_gss: BH069944	61.50	132.31	401.20	654	I BH069944 RPCI-24-361N6 TV RP
gb_gss: A2705220	62.00	129.65	564.17	932	I A2705220 Tetraodon nigroviridis	gb_est1:AL655783	61.50	132.26	403.64	657	I AL655783 XGC-neurul
gb_gss: A2705220	62.00	129.56	571.24	968	I AQ688092 602364655F1 NIH_MGC.9C	gb_est1:AL643421	61.50	132.25	404.45	658	I AL643421 AL643421 XGC-neurul
gb_est2:BF2252015	62.00	129.48	576.75	968	I BG252015 602364655F1 NIH_MGC.9C	gb_gss: A2981374	61.50	132.10	410.15	661	I A2981374 2M0255H14R Mouse 10
gb_gss: A2705220	62.00	128.86	624.50	1028	I AL174573 Tetraodon nigroviridis	gb_est2:BF160409	61.50	132.12	410.97	666	I BF160409 EPT562932 PyBS Plas
gb_gss: A2705220	62.00	128.67	639.81	1047	I BF108009 601824058R1 NIH_MGC.7	gb_est1:AL650068	61.50	132.11	411.78	667	I AL650068 AL650068 XGC-gastru
gb_est2:BF399383	62.00	128.35	666.61	1080	I BE735786 6013050509F1 NIH_MGC.3	gb_est1:BB630001	61.50	132.08	413.42	669	I BB630001 BB630001 RIKEN full
gb_est2:BF399383	62.00	128.04	713.54	1137	I BG390732 602416446F1 NIH_MGC.9	gb_est2:BI295022	61.50	132.08	413.42	669	I BI295022 UI-R-DK0-ccc-h-03-0
gb_est1:BM464281	62.00	127.82		1137	I BM464281 AGENCOURT_6439015 NIH	gb_est1:AL638209	61.50	131.88	424.08	682	I AL638209 XGC-neurul



gb_gss: BH036363	61.50	131.74	431.49	691	BH036363	RPCI-24-295A15.TV RPCI	gb_gss: A2843321	61.00	131.75	431.26	614	AZ843321	2M0142G08F Mouse 10
gb_est2: BG593493	61.50	131.71	433.15	693	BG593493	EST492171 CSTS Solanum	gb_gss: AZ007704	61.00	131.73	432.18	615	AZ007704	RPCI-23-246123.TV R
gb_gss: A2557685	61.50	131.68	434.80	695	A2557685	RPCI-23-217EG.TV RPCI	gb_gss: AZ010519	61.00	131.61	438.70	622	AZ010519	RPCI-24-222K19.TJ R
gb_est2: B1697769	61.50	131.65	436.46	697	B1697769	603346683F1 NCI_CGAP_M	gb_est2: BE356398	61.00	131.56	441.31	625	BE356398	DG1_124_D03_g1 A0043
gb_gss: BH072341	61.50	131.60	439.77	701	BH072341	RPCI-24-286G323.TV RPCI	gb_est2: BE4406903	61.00	131.52	444.31	628	BE4406903	SP49606.y1 Gm-cl043
gb_est1: BB646819	61.50	131.54	443.10	705	BB646819	BB646819 RIKEN full1-le	gb_est2: BMA367112	61.00	131.42	449.93	634	BMA367112	VVA008F09_530075.TJB R
gb_est2: B1660549	61.50	131.46	447.26	710	B1660549	603303633F1 NCI_CGAP_M	gb_gss: AZ908718	61.00	131.40	450.87	635	AZ908718	RPCI-24-109B5.TJB R
gb_gss: BH0605870	61.50	131.42	449.76	713	BH0605870	RPCI-24-361B21.TJ RPCI	gb_hic: AK007129	61.00	131.35	453.69	638	AK007129	Mus musculus adult
gb_est1: BH344418	61.50	131.32	455.61	720	BH344418	CH230-52A1F1.TV CHORI-2	gb_gss: BH3610264	61.00	131.27	459.34	643	BH3610264	CH230-95M6.TJ CHORI
gb_est2: B1697944	61.50	131.21	462.31	728	B1697944	603340481F1 NCI_CGAP_M	gb_est1: AW309264	61.00	131.26	459.34	644	AW309264	CH29202.x1 Gm-cl028
gb_gss: A2616461	61.50	131.19	463.16	729	A2616461	BJ146461 unpublished c	gb_est1: AW5081229	61.00	130.96	477.35	663	AW5081229	AV911344 K. Sato un
gb_gss: A2616461	61.50	131.16	464.84	731	A2616461	BJ146461 unpublished c	gb_est1: AW5081229	61.00	130.96	477.35	663	AW5081229	AV911344 K. Sato un
gb_gss: BH222880	61.50	131.02	473.27	741	BH222880	CH230-5318.TV CHORI-23	gb_gss: A2367264	61.00	130.86	483.07	669	A2367264	1M0116A21R Mouse 10
gb_est1: AL046546	61.50	131.00	474.96	743	AL046546	DFEP2434Q1725.r1 434 c	gb_est1: AL046546	61.00	130.79	487.86	674	AL046546	DFEP2434Q1725.r1 434 c
gb_gss: BH33746	61.50	130.94	478.34	747	BH33746	EST566269 PYS Brassic	gb_est1: BH117486	61.00	130.79	487.86	674	BH117486	BH117486 RIKEN full
gb_gss: BH434800	61.50	130.83	485.13	755	BH434800	BOGDY88TR BOGD Brassic	gb_est1: BH117486	61.00	130.79	487.86	674	BH117486	BH117486 RIKEN full
gb_gss: A2817494	61.50	130.75	490.23	761	A2817494	2M0087D04F Mouse 10kb	gb_est1: A1225192	61.00	130.68	494.57	681	A1225192	U332604.y1 Soares m
gb_gss: CN507210	61.50	130.63	497.92	770	CN507210	CLONE BA008A007B12 of 1	gb_gss: BH430497	61.00	130.41	511.93	699	BH430497	BGMK34TR BGMK Bras
gb_gss: BH338802	61.50	130.59	500.48	773	BH338802	CH230-130G15.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_est1: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_est1: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF
gb_gss: BH297017	61.50	130.51	505.63	779	BH297017	CH230-57EA.TV CHORI-23	gb_gss: A2404589	61.00	130.37	514.84	702	A2404589	CIT-HSP-2387F21.TF



gb_est1.AV536967	+	60.50	133.51	343.94	460	AV536967	Arabidopsis
gb_est2.BF014594	-	60.50	133.44	346.91	463	BFI014594	RCO-ET0129-060201-0322
gb_est2.BF565050	-	60.50	133.36	350.88	467	BF565050	UI-R-B01-ajm-g-06-0-01
gb_gss.AZ314727	+	60.50	133.25	355.86	472	AZ314727	tt0210J10R Mouse 10kb
gb_est2.BI594837	+	60.50	133.18	358.85	475	BI594837	As_tgz-44G11_SKPL Asca
gb_gss.BH285313	+	60.50	133.18	358.85	475	BH285313	CH230-137H9_TV CHORI-2
gb_gss.BH285319	+	60.50	133.12	361.85	478	BH285319	476.d1046b03.s1 Saccha
gb_est1.AW639704	+	60.50	132.88	372.91	489	AW639704	b16ga05.w1 Blackshear
gb_est1.AV533558	+	60.50	132.82	375.94	492	AV533558	AV533558 Arabidopsis t
gb_gss.BH309875	+	60.50	132.69	382.01	498	BH309875	CH230-110B15_TV CHORI-1
gb_est2.BG226448	+	60.50	132.65	384.04	500	BG226448	kp89c08.y1 TB9N5TM-SSR
gb_est2.BF382653	+	60.50	132.61	386.08	502	BF382653	601816862F1 NIH_MGC-56
gb_gss.BH305066	+	60.50	132.59	387.09	503	BH305066	CH230-192J10_TV CHORI-1
gb_est2.BG39134	+	60.50	132.53	390.15	506	BG39134	LD44332.3prime LD Dros
gb_est2.BE334327	+	60.50	132.53	390.15	506	BE334327	u41d08.y1 Soares_NMBE
gb_est1.AA051644	+	60.50	132.51	391.17	507	AA051644	m155d12.l1 Soares mouse
gb_gss.AZ242448	+	60.50	132.41	396.28	512	AZ242448	RPCI-23-87B17_TV RPCI-1
gb_gss.AV277905	+	60.50	132.39	397.30	513	AZ277905	RPCI-24-98B7_TV RPCI-2
gb_est1.AV534574	+	60.50	132.33	400.38	516	AV534574	AV534574 Arabidopsis
gb_gss.AQ416534	+	60.50	132.33	400.38	516	AQ416534	RPCI-11-173N7_TV RPCI-1
gb_gss.AZ244299	+	60.50	132.31	401.41	517	AZ244299	RPCI-23-56H15_TV RPCI-1
gb_gss.AQ930236	+	60.50	132.29	402.43	518	AQ930236	RPCI-23-265L11_TV RPCI-1
gb_est1.AV587681	+	60.50	132.27	403.46	519	AV587681	vt14e07.y1 Barstead m
gb_gss.AZ373382	+	60.50	132.09	412.74	528	AZ373382	1M0425F07R Mouse 10kb
gb_est1.AA098141	+	60.50	132.01	416.88	532	AA098141	mm84d07.r1 Strategene
gb_gss.BH293479	+	60.50	131.76	430.41	545	BH293479	CH230-172M7_TV CHORI-2
gb_gss.AZ095411	+	60.50	131.69	434.59	549	AZ095411	RPCI-23-461F23_TV RPCI-2
gb_gss.BH297305	+	60.50	131.67	435.64	550	BH297305	CH230-123L1_TV CHORI-2
gb_gss.BH045299	+	60.50	131.67	435.64	558	BH045299	RPCI-24-242G12_TV RPCI-1
gb_gss.BH182333	+	60.50	131.43	449.31	563	BH182333	021_C_19-21 SmBAIC1 SC
gb_gss.CNS07NKO	+	60.50	131.43	449.31	563	CNS07NKO	T7 end of clone 021ABE
gb_gss.AZ445631	+	60.50	131.39	451.42	565	AZ445631	1M0241J05R Mouse 10kb
gb_est1.AM180298	+	60.50	131.36	453.54	567	AM180298	MGA0392F MGA Library M
gb_gss.AZ116211	+	60.50	131.28	457.78	571	AZ116211	RPCI-23-44SNA_TV RPCI-1
gb_gss.AZ959594	+	60.50	131.26	458.84	572	AZ959594	2M0282F04F Mouse 10kb
gb_est1.AV591467	+	60.50	131.25	459.90	573	AV591467	vt14e07.x1 Barstead m
gb_est1.AV593980	+	60.50	131.23	460.96	574	AV593980	bl87d02.w1 Blackshear
gb_est2.BF473283	+	60.50	131.21	462.02	575	BF473283	333080 MARC 1P1G Sus
gb_gss.AQ930719	+	60.50	131.19	463.08	576	AQ930719	maa39pb06.x1 NCI_CGAP
gb_gss.AQ930719	+	60.50	131.19	463.08	576	AQ930719	RPCI-23-265N13_TV RPCI-1
gb_est1.AW527212	+	60.50	131.17	464.15	577	AW527212	UI-R-B01-ajm-g-06-0-01
gb_gss.BH309562	+	60.50	131.17	464.15	577	BH309562	CH230-98K17_TV CHORI-2
gb_est1.AV599239	+	60.50	131.14	466.28	578	AV599239	EST309739 L. pennellii
gb_est2.BF466029	+	60.50	131.03	472.68	585	BF466029	UI-M-CG0P-bqs-f-12-0-11
gb_gss.AZ2936307	+	60.50	131.03	472.68	585	AZ2936307	RPCI-23-108J18_TV RPCI-1
gb_est1.AW687503	+	60.50	131.03	472.68	589	AW687503	NFO10C09RT1F1069 Devel
gb_gss.BH083652	+	60.50	130.96	476.96	589	BH083652	RPCI-24-293F12_TV RPCI-1
gb_est1.AU12676	+	60.50	130.94	476.96	590	AU12676	AU127676 N2R2P2 Homo s
gb_gss.BH0322757	+	60.50	130.91	480.17	592	BH0322757	RPCI-24-368P22_TV RPCI-2
gb_gss.BH290500	+	60.50	130.79	487.70	599	BH290500	CH230-30B13_TV CHORI-2
gb_gss.BH295366	+	60.50	130.77	488.78	600	BH295366	CH230-92122.TJB CHORI-2
gb_est1.BB622229	+	60.50	130.67	495.25	606	BB622229	BB622229 RIKEN full-le
gb_gss.BH295827	+	60.50	130.62	498.50	609	BH295827	CH230-57B14_TV CHORI-1
gb_gss.AZ875864	+	60.50	130.43	510.44	620	AZ875864	2M0190D12R Mouse 10kb
gb_gss.AZ241420	+	60.50	130.35	515.89	625	AZ241420	1M0199K09R Mouse 10kb
gb_gss.AA550043	+	60.50	130.14	530.13	638	AA550043	1128m3 gmbpFHB3.1, G.
gb_gss.AZ659323	+	60.50	130.11	532.33	640	AZ659323	1M0536F09F Mouse 10kb
gb_est2.BF632340	+	60.50	130.09	533.43	641	BF632340	NF027F08DT1F1064 Drou
gb_gss.AQ639560	+	60.50	130.04	536.74	644	AQ639560	927P1-13A3_TV 927P1 Th
gb_gss.AZ660572	+	60.50	130.01	538.94	646	AZ660572	1M0538C14R Mouse 10kb
gb_gss.AZ2720091	+	60.50	129.96	542.25	649	AZ2720091	RPCI-24-144B4_TV RPCI-1
gb_gss.AZ937747	+	60.50	129.95	543.36	650	AZ937747	2M0196K01F Mouse 10kb
gb_gss.AZ522802	+	60.50	129.91	545.57	652	AZ522802	211PB005 Pb MBN #21 Pl
gb_est2.BF297653	+	60.50	129.88	547.79	654	BF297653	054PbD11 Pb cdNA #17,
gb_gss.AZ41087	+	60.50	129.87	548.90	655	AZ41087	1M0266G13R Mouse 10kb
gb_est2.BM168710	+	60.50	129.77	555.56	661	BM168710	EST571233 PVBS Plasmid
gb_est2.BF636652	+	60.50	129.77	555.56	661	BF636652	NFO67E11ST1F1086 Devel
gb_gss.AZ701970	+	60.50	129.76	556.67	662	AZ701970	RPCI-23-22SA11_TV RPCI-1
gb_gss.BH278731	+	60.50	129.74	557.78	663	BH278731	CH230-178J23_TV CHORI-1
gb_gss.AZ336664	+	60.50	129.73	558.89	664	AZ336664	1M0067J04F Mouse 10kb
gb_gss.AZ860284	+	60.50	129.71	560.01	665	AZ860284	2M0166J03F Mouse 10kb
gb_gss.BH425823	+	60.50	129.70	561.12	666	BH425823	BOGFR64TR BOGF Brassid
gb_gss.AZ450856	+	60.50	129.60	567.82	672	AZ450856	1M0249J118R Mouse 10kb

gb_gss:BH012925	+	60.50	129.59	568.93	673	! BH012925	TDGAQ10TH cTOG Lyco
gb_est1.BB642270	-	60.50	129.50	575.65	679	! BB642270	BB642270 RIKEN full
gb_gss:AZ860245	-	60.50	129.50	575.65	679	! AZ860245	2M0166H05F Mouse 10
gb_gss:AG135546	-	60.50	129.50	575.65	679	! AG135546	Pan troglodytes DNA
gb_gss:AZ273514	-	60.50	129.42	581.27	684	! AZ273514	RPCI-23-147C1.TV RP
gb_gss:AZ273514	-	60.50	129.36	585.77	688	! AZ273514	2M015H05F Mouse 10
gb_gss:AZ971248	-	60.50	129.21	597.05	698	! AZ971248	2M0244K11R Mouse 10
gb_est1.AL584098	-	60.50	129.15	601.58	702	! AL584098	AL584098 StrataGene
gb_gss:BH32602	+	60.50	129.15	601.58	702	! BH32602	CH230-20IN11.TV CHO
gb_gss:BH125143	+	60.50	129.14	602.72	703	! BH125143	RPCI-24-306FA.TV RP
gb_gss:BH038696	-	60.50	129.08	607.26	707	! BH038696	RPCI-24-35907.TJB RP
gb_gss:BH027265	-	60.50	129.02	611.81	711	! BH027265	RPCI-24-25943.TV RP
gb_gss:BH071616	-	60.50	128.96	616.36	715	! BH071616	RPCI-24-39A1A3.TV R
gb_gss:AZ959614	-	60.50	128.88	623.21	721	! AZ959614	2M0227B19F Mouse 10
gb_gss:AZ657989	-	60.50	128.83	626.64	724	! AZ657989	1M0534G03R Mouse 10
gb_gss:BH269504	-	60.50	128.69	638.12	733	! BH269504	CH230-176K3.TJ CHOR
gb_gss:AZ247659	-	60.50	128.65	641.57	737	! AZ247659	RPCI-23-57A12.TJ RP
gb_gss:BH201496	+	60.50	128.57	648.49	743	! BH201496	CH230-43P22.TV CHOR
gb_gss:BH591159	-	60.50	128.51	653.11	747	! BH591159	BOH0036TR BOHK Brs10
gb_gss:AZ315163	-	60.50	128.48	655.42	749	! AZ315163	1M0032H18F Mouse 10
gb_gss:BH263567	-	60.50	128.46	657.74	751	! BH263567	CH230-17H20.TVC CH
gb_gss:AZ655860	-	60.50	128.36	665.86	758	! AZ655860	1M0531M02F Mouse 10
gb_gss:BH360007	+	60.50	128.35	667.02	759	! BH360007	CH230-68R.TJ CHORI-
gb_est2.BM169065	-	60.50	128.32	669.35	761	! BM169065	EST571588 PYBS Plas
gb_gss:BH332081	-	60.50	128.29	671.68	763	! BH332081	CH230-220K16.TV CHO
gb_gss:BH056403	-	60.50	128.25	675.17	766	! BH056403	RPCI-24-33738.TV RP
gb_gss:BH523238	-	60.50	128.20	679.84	770	! BH523238	BOGLH24TR BOGL Brasi
gb_gss:BH256631	-	60.50	128.09	689.20	778	! BH256631	CH230-286A19.TJ CHO
gb_gss:AQ779910	-	60.50	127.99	698.59	786	! AQ779910	HS_3104.AL.F10.T7C
gb_gss:BH535948	-	60.50	127.91	705.65	792	! BH535948	BOGGQ13TR BOGG Brass
gb_gss:BH077860	+	60.50	127.78	717.46	802	! BH077860	CH230-17H9.TJ CHORR
gb_gss:BH073136	+	60.50	127.72	723.38	807	! BH073136	RPCI-24-303D22.TV R
gb_est1.AU124072	-	60.50	127.69	725.75	809	! AU124072	AU124072 N2R2M2 Hom
gb_gss:BH329821	-	60.50	127.54	740.02	821	! BH329821	CH230-465F.TV CHOR
gb_est2.B1952645	-	60.50	127.41	751.97	831	! B1952645	HVSMem0007D12f Hordr
gb_est2.B1907810	-	60.50	127.40	753.16	832	! B1907810	Q0086072F1 NIH_MGC
gb_gss:AZ530264	+	60.50	127.39	754.36	833	! AZ530264	ENTD432TR Entamoeba
gb_gss:AZ533116	-	60.50	127.23	769.97	846	! AZ533116	ENTC033TR Entamoeba
gb_gss:BH115914	-	60.50	127.14	778.41	853	! BH115914	ENTQ009TR Entamoeba
gb_gss:AZ285239	-	60.50	127.04	789.29	862	! AZ285239	RPCI-23-125K10.TJ R
gb_gss:CN507436	+	60.50	126.87	806.29	876	! CN507436	clone BA0A80A030H10
gb_gss:BH157964	-	60.50	126.60	834.41	899	! BH157964	ENTRP80TR Entamoeba
gb_gss:AZ668140	-	60.50	126.54	840.55	904	! AZ668140	ENTK212TR Entamoeba
gb_gss:AZ689749	-	60.50	126.42	854.11	915	! AZ689749	ENTHL54TR Entamoeba
gb_gss:CN5021R8	-	60.50	126.03	897.58	950	! CN5021R8	TETraodon nigrovir
gb_gss:BH157952	-	60.50	126.01	900.08	952	! BH157952	TETraodon nigrovir
gb_gss:CN502PEY	-	60.50	125.78	927.70	974	! CN502PEY	TETraodon nigrovir
gb_gss:CN507053	+	60.50	125.57	952.98	994	! CN507053	TETraodon nigrovir
gb_est2.BE562722	-	60.50	125.50	960.59	1000	! BE562722	601336003F1 NIH_MGC
gb_est1.AV171553	+	60.50	125.26	991.20	1024	! AV171553	AV171553 Mus muscu
gb_gss:CN5045V8	+	60.50	124.92	1.0e+03	1058	! CN5045V8	TETraodon melano
gb_gss:CN5012M0	+	60.50	124.52	1.1e+03	1100	! CN5012M0	AL101756 Drosophila
gb_est2.B1552505	+	60.50	124.23	1.1e+03	1131	! B1552505	603194416F1 NIH_MGC
gb_hcc:AK006980	-	60.50	118.25	2.4e+03	2021	! AK006980	Mus musculus adult
gb_est1.AV808026	-	60.50	116.36	3.1e+03	2426	! AV808026	AV808026 Mus muscu
gb_est1.AV808011	-	60.50	135.57	251.77	323	! AV808011	AV808011 Nori Satoh
gb_est1.AV860368	-	60.00	135.94	255.90	327	! AV860368	AV860368 Nori Satoh
gb_est1.AV860368	-	60.00	135.82	255.90	327	! AV860368	AV860368 Nori Satoh
gb_est2.N55783	-	60.00	135.75	257.98	329	! N55783	J2732F Human fetal
gb_gss:BH345913	+	60.00	135.75	257.98	329	! BH345913	CH230-120A20.TJ CHO
gb_est1.AV808026	-	60.00	135.60	263.12	334	! AV808026	AV808026 Nori Satoh
gb_est1.AV808011	-	60.00	135.57	264.22	335	! AV808011	AV808011 Nori Satoh
gb_est1.AV860365	-	60.00	135.48	267.35	338	! AV860365	AV860365 Nori Satoh
gb_est1.AV868149	-	60.00	135.35	271.54	342	! AV868149	AV868149 Nori Satoh
gb_est1.AV887366	-	60.00	135.35	271.54	342	! AV887366	AV887366 Nori Satoh
gb_est1.AV840353	-	60.00	135.32	272.59	343	! AV840353	AV840353 Nori Satoh
gb_est1.AV859404	-	60.00	135.32	272.59	343	! AV859404	AV859404 Nori Satoh
gb_est1.AV875438	-	60.00	135.29	273.65	344	! AV875438	AV875438 Nori Satoh
gb_est1.AV863341	-	60.00	135.26	274.70	345	! AV863341	AV863341 Nori Satoh
gb_est1.AV860937	-	60.00	135.18	277.86	348	! AV860937	AV860937 Nori Satoh
gb_est1.AV802728	-	60.00	135.12	279.98	350	! AV802728	AV802728 Nori Satoh
gb_est1.AV846147	-	60.00	134.88	288.47	358	! AV846147	AV846147 Nori Satoh
gb_est1.AV679977	-	60.00	134.57	300.25	369	! AV679977	AV679977 Nori Satoh
gb_est1.AA470987	+	60.00	134.24	313.24	381	! AA470987	ve60812_r1 Beddingt







```

77 ATGACACTCTTCAATAGTCCAAATAATATAAAGTACCTCGGTGTGAC 126
80 .IleSerAlaAsnVallys.....TyrAspMetAsnPhelYsLysLeuL 94
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
127 TTAAACCAAGCAAGTGAAGATCTATATGACAGAAGACTTCAAGTCTCTGA 176
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
94 euArgPheMetGluAspAspPheIle 102
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
177 AGAAGAAGAAATGAAGAAGACCTCATA 202

```

seq\_name: gb\_gss:BH332419

seq\_documentation\_block:  
LOCUS BH332419 660 bp DNA linear GSS 03-DEC-2001  
DEFINITION CH230-189M18-TV CHORI-230 Segment-1 Rattus norvegicus genomic clone  
ACCESSION BH332419  
VERSION BH332419.1 GI:17263133  
KEYWORDS GSS.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 660)  
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn  
A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
Jong,P. and Fraser,C.M

TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment

JOURNAL Unpublished (1999)  
COMMENT Other\_GSSs: CH230-189M18.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources.  
page: http://www.chori.org/bacpac/orering\_information.htm). BAC end  
plate: 189 row: M column: 18  
Seq primer: T7  
Class: BAC ends.

FEATURES Location/Qualifiers  
1..660  
/organism="Rattus norvegicus"  
/strain="BN/SSNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-189M18"  
/clone\_lib="CHORI-230 Segment 1"  
/sex="Female"  
/cell\_type="Brain"  
/note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
Pieter de Jong"

BASE COUNT 268 a 114 c 117 g 161 t

ORIGIN

alignment\_scores:  
Quality: 76.50 Length: 41  
Ratio: 2.638 Gaps: 2  
Percent Similarity: 70.732 Percent Identity: 43.902

alignment\_block:  
US-09-528-682-3 x BH332419

Align seg 1/1 to: BH332419 from: 1 to: 660

```

64 MetSerThrPheGluGlnValProAsnAsnLysGluPhelYsGlyVal.. 79
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
337 ATGACACCTTCAATAATCCCAATAATATAAATATCTTGGTGTGAC 386
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
80 .IleSerAlaAsnVallys.....TyrAspMetAsnPhelYsLysLeuL 94
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
387 TTTAAGTACCAAGTAAAAAATTGTATGACAGAAGAACTTCAAGTCTCTAA 436
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
94 euArgPheMetGluAspAspPhe 101
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
437 AGAAGAAGAAATGAGGAAGATTTC 459

```

seq\_name: gb\_est2:BI316490

seq\_documentation\_block:  
LOCUS BI316490 540 bp mRNA linear EST 29-NOV-2001  
DEFINITION saf03h08.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl065-4576 5' similar to TR:Q92W71 Q92W71 F14B2.18 PROTEIN.  
; mRNA sequence.

ACCESSION BI316490  
VERSION BI316490.1 GI:14990817  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 540)  
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna  
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
R., Waterston,R. and Willson,R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 418.

FEATURES Location/Qualifiers  
1..540

source

/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-4576"  
/clone\_lib="Gm-cl065"  
/tissue\_type="germinating shoots"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from mRNA isolated  
germinating shoots of the cultivar Williams. The seeds  
were allowed to germinate for 24 hours prior to being  
cold stressed for 2 days at 4C. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by XhoI digestion. The cDNA fragments were  
directionally cloned into the EcoRI-XhoI restriction site  
of the pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This library  
was constructed in the laboratory of Dr. Randy  
Shoemaker."

BASE COUNT 122 a 133 c 104 g 181 t  
ORIGIN



```
alignment_scores:
  Quality: 76.00      Length: 99
  Ratio: 1.357       Gaps: 6
  Percent Similarity: 56.566      Percent Identity: 29.293

alignment_block:
  US-09-528-682-3 x BI316490/rev ..

  Align seg 1/1 to reverse of: BI316490 from: 1 to: 540

14  lIeGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :||||:||||| |||||:||||: ||: ||: |||
462 CTGAAGCAACTAGACACAATCACCACGACCAAAAAGACGTTGAGTG 413

30  lProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
   ||| |||||:|||||:||||: |||
412 CACGTTGCTTGGACGATATCTACTTTTGTGGCATCATAAAGCCCTTGAT 363

43  ..SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
   ||| ||||| ||||| ||||| ||||| |||||
362 ATAGTTTGGCGGAGG.....GGAAACCAATTGGATGCGCAAC 322

59  GlnSer...AsnPhePro.....MetSerTh 66
   :||: |||||:|||||:||||: |||
321 ACTGCATTCATATCCGGTTCTTCATGATATATCCTCCACACATATCAGG 272

66  rPheGluGlnValProAsnAsn.....LysGluPheLysGlyValI 80
   :||||:||||: ||| ||||| ||||| |||||
271 GTATGAAGAAGCTGATATAAATGAGTCCAGAAAGACATTCGGGT.... 226

80  lSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArg 95
   :||: ||||| ||||| ||||| ||||| |||||
225 .....GAACTCGATGTGATCTGCAGCTGAGCGGACGATTTCGA 187

seq_name: gb_est2:BM270616

seq_documentation_block:
  LOCUS      BM270616          554 bp      mRNA      linear      EST 20-DEC-2001
  DEFINITION  Sak15d11.y1 Gm-c1075 Glycine max cDNA clone SOYBEAN CLONE ID:
  Gm-c1075-3237 5' similar to TR:Q92W71 Q92W71 F14B2.18 PROTEIN. ;
  mRNA sequence.
  BM270616.1 GI:17963872
  VERSION    BM270616.1
  KEYWORDS   soybean.
  SOURCE     Glycine max
  ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
  1 (bases 1 to 554)
  Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
  ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
  ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  ,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  call: (800)-533-4363 or contact: ccu@resgen.com web site:
  www.resgen.com
  Seq primer: -40RP from Gibco
```

```
High quality sequence stop: 425.
Location/Qualifiers
  1..554
  /organism="Glycine max"
  /db_xref="taxon:3847"
  /clone="SOYBEAN CLONE ID: Gm-c1075-3237"
  /clone_lib="Gm-c1075"
  /tissue_type="differentiating somatic embryos cultured on
  MSMGAC"
  /lab_host="DH10B"
  /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
  XhoI; The cDNA library was constructed from mRNA isolated
  from differentiating somatic embryos cultured on MSMGAC.
  The library was prepared using the Stratagene pBluescript
  II SK(+) library construction kit. Complementary DNA was
  synthesized from mRNA using a primer consisting of a
  poly(dT) sequence with an XhoI restriction site. EcoRI
  adaptors were ligated to the blunt-ended cDNA fragments
  followed by XhoI digestion. The cDNA fragments were
  directionally cloned into the EcoRI-XhoI restriction site
  of the pBluescript vector. The ligated cDNA fragments
  were transformed into E.coli ElectroMax DH10B host cells.
  Tissue culture and library construction were performed by
  Francoise Thibaud-Nissen and Anu Khana (Lila Vodkin lab,
  University of Illinois)."
  BASE COUNT      127 a      135 c      106 g      186 t
  ORIGIN

alignment_scores:
  Quality: 76.00      Length: 99
  Ratio: 1.357       Gaps: 6
  Percent Similarity: 56.566      Percent Identity: 29.293

alignment_block:
  US-09-528-682-3 x BM270616/rev ..

  Align seg 1/1 to reverse of: BM270616 from: 1 to: 554

14  lIeGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :||||:||||| |||||:||||: ||: ||: |||
461 CTGAAGCAACTAGACACAATCACCACGACCAAAAAGACGTTGAGTG 412

30  lProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
   ||| |||||:|||||:||||: |||
411 CACGTTGCTTGGACGATATCTACTTTTGTGGCATCATAAAGCCCTTGAT 362

43  ..SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
   ||| ||||| ||||| ||||| ||||| |||||
361 ATAGTTTGGCGGAGG.....GGAAACCAATTTGGAATGGCAAC 321

59  GlnSer...AsnPhePro.....MetSerTh 66
   :||: |||||:|||||:||||: |||
320 ACTGCATTCATATCCGGTTCTTCATGATATATCCTCCACACATATCAGG 271

66  rPheGluGlnValProAsnAsn.....LysGluPheLysGlyValI 80
   :||||:||||: ||| ||||| ||||| ||||| |||||
270 GTATGAAGAAGCTGATCTAAATCGAGTCCAGAAAGACGATTCGGGT.... 225

80  lSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArg 95
   :||: ||||| ||||| ||||| ||||| |||||
224 .....GAACTCGATGTGATCTGCAGCTGAGCGGACGATTTCGA 186

seq_name: gb_est1:AL636005

seq_documentation_block:
  LOCUS      AL636005          723 bp      mRNA      linear      EST 12-DEC-2001
  DEFINITION  AL636005 XGC-neurula silurana tropicalis cDNA clone tNeu015014 5',
  mRNA sequence.
  ACCESSION  AL636005
  VERSION    AL636005.1
  KEYWORDS   EST.
  SOURCE     western clawed frog.
```







```

|||||||
270 GTAATTTTCACTACCGACTTAATGTTTGTAAACATATCTCTCAGAAATGTA 319
|||||||
67 PheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAlaAs 83
|||||||
320 TTTTATATTGATCTCTCAAAAATGAATACCAGATTA.....AA 360
|||||||
83 nValLysTyrAspMetAsnPhe 90
|||||||
361 TTTAGATTATGACTATTAATCTAT 382
|||||||

seq_name: gb_gss:BH266151

seq_documentation_block:
LOCUS BH266151 322 bp DNA linear GSS 30-NOV-2001
DEFINITION CH230-55L4.TV CHORI-230 segment 1 Rattus norvegicus genomic clone
ACCESSION BH266151
VERSION BH266151.1 GI:17176966
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 322)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other_GSSs: CH230-55L4.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 55 row: L column: 4
Seq primer: T7
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..322
/organism="Rattus norvegicus"
/strain="BN/SsNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-55L4"
/clone_lib="CHORI-230 Segment 1"
/sex="female"
/cell_type="Brain"
/notes="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SsNhsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT 80 a 56 c 55 g 131 t
ORIGIN

alignment_scores:
Quality: 75.00 Length: 51
Ratio: 2.273 Gaps: 3
Percent Similarity: 64.706 Percent Identity: 39.216

alignment_block:
US-09-528-682-3 x BH266151/rev ..
Align seg 1/1 to reverse of: BH266151 from: 1 to: 322

|||||||
64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79
|||||||
290 ATGACACCCCTTCAAAATAGTCCCAATAATAATAATACCTTGGTGTGAG 241
|||||||
80 IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94
|||||||
240 TTTAACCAAGGAAGTGAAGATCTGTATGATAAGAACTTCAAGCCTCTGA 191
|||||||
94 euArgPheMetGluAspAspPhe.....IleGlyValHis 105
|||||||
190 AGAAGAAATTGAAGAGATCTCAGAAGGTGGAAAGATCTCCATGCTCAT 141
|||||||
106 Gly 106
|||||||
140 GGA 138

seq_name: gb_est1:A1134063

seq_documentation_block:
LOCUS A1134063 583 bp mRNA linear EST 19-APR-2001
DEFINITION GH11413.5prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH11413 5prime similar to X61209: Top2
FBgn0003732 PID:98711 SWISS-PROT:P15348, mRNA sequence.
ACCESSION A1134063
VERSION A1134063.1 GI:3626621
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 583)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HHMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd. Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic sequence DS05231
Plate: 114 row: B column: 1
High quality sequence stop: 534.
FEATURES
source Location/Qualifiers
1..583
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH11413"
/clone_lib="GH Drosophila melanogaster head pOT2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/notes="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. plasmid cDNA library."
BASE COUNT 150 a 140 c 124 g 169 t
ORIGIN

alignment_scores:
Quality: 75.00 Length: 118
Ratio: 1.230 Gaps: 5
Percent Similarity: 51.695 Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x A1134063/rev ..
Align seg 1/1 to reverse of: A1134063 from: 1 to: 583

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlySe 27

```







TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Professionalism	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 458.

Location/Qualifiers  
1. .471

BASE COUNT  
ORIGIN

```
alignment_block:
  US-00-528-682-3
```

**190** GAAAGATGGGATCGCATATAA...TTTCTCAGCTCAACACATAA 236

28 nGlu...ValProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerA 44  
 |||| :: |||||| |||::: ||| |||:  
 237 CGAGGATGGACTGGATGGTTTATGCTCACAACTGACTTGTCACTT 286

44 laGluGlyClyMetGlnAspGly.....Asp 52  
:::| | | | |  
287 CTGAACCTGCACCTCCGATCGGTGAGGGCGGACGATGACACTGCAGAT 336

53 LeuPheGlyThrValAsnGln...SerAsnPhePromet.....Se 65  
||| |||||:: ::|||::: ||  
337 TGGGTTGGCCCTGTAATGAAAGACCAAAATTTCTCTCTCGGCAAAATC 386

65 rThr-Phe-Glu-Gln-Val-Pro-Asn-Lys-Glu-Phe-Lys-Gly-Val-Ile-Ser-A 82  
||||| :||| |::||| :::  
387 AACGAAACAGCAG.....CGTTTATTAACGGCTACTG 421

82 laAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPheMetGlu 98  
|||::: :::: ||||| |||||::: :::  
422 CATCAGCAGCTCTCGAG.....CGAAACTTGAAAGATTTATTTCGA 462

422 CATCAGCAGCTCTGCAG.....CGGAACTTGAAAGATTATTCTGA 462

seq\_name: gb\_est1:AA333548

```
seq_documentation_block:
  r_code: 1822540
```

LOCUS	294 bp	linear	EST 21-APR-1997
RA333548	294 bp	linear	EST 21-APR-1997
DEFINITION	EST37697	Embryo, 8 week I Homo sapiens	cdna 5' end similar to
		aggrecan 1, mRNA sequence.	

ACCESSION  
AA333548  
VERSION  
AA333548.1  
KEYWORDS  
EST.  
SOURCE  
human.

SOURCE	ORGANISM	Human.	Homo sapiens
...	...	...	...

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 294)

REFERENCE  
ADAMS, M.D. to 294)  
C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodok, A., Gnehm, C.J.D., Hanna, M.C., Hedblom, E., Hinkie, P.P.S.-Dr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marnaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, L., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hwang, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Hing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

MEDLINE	90026280	Other_ESTs: THC166485
COMMENT		

Contact: Kerlavage, AR  
Bioinformatics

**PROBIOGENOMICS**  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423

Email: [arkerlav@tigr.org](mailto:arkerlav@tigr.org)  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Accession: [U15282](#)

```

FEATURES
  source      Location/Qualifiers
seq primer: M13 reverse.

```

```

1. .294
organism="Homo sapiens"
db_xref="ATCC (inhost):135261"
db_xref="taxon:9606"
clone_lib="Embryo, 8 week I"
dev_stage="embryo, 8 wks"
note="Organ: Embryo, 8 weeks;

```

BASE COUNT	Site_1: EcoRI; Site_2: XhoI <sup>a</sup>	
69 a	62 c	86 g
74 t		
3 others		

alignment_scores:			
Quality:	73.50	Length:	70
Ratio:	1.670	Gaps:	2
percent Similarity:	62.857	Percent Identity:	34.286

alignment\_block:  
US-09-528-682-3

US-09-528-682-3 x AA333548



Align seg 1/1 to: AA333548 from: 1 to: 294

```

14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :::::||||| :::::||||| :::::||||| :::::|||||
44 GTTGAAGTGGCCCTACTACATTTAAAGAAGAAGAGCTTAGGGTCTGT 93
   | :::::||||| :::::||||| :::::||||| :::::|||||
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly 47
   | :::::||||| :::::||||| :::::||||| :::::|||||
94 GGAACACAGTGGCCCTCCCTCCGGAGAGCAGATCTGTCCAGCAATCTG 143
   | :::::||||| :::::||||| :::::||||| :::::|||||
47 lYMetGlnAsp.....GlyAspLeuPheGlyThrValAsnGlnSerAsn 61
   ||||| ||| ||||| :::::||||| :::::||||| :::::|||||
144 GGATGGTGGATGTCACGTGGACAGATTTTNGAAGACAGTCGATTCAGTGG 193
   | :::::||||| :::::||||| :::::||||| :::::|||||
62 PheProMetSerThrPheGluGlnValProAsnAsnLysGluPheLysGl 78
   ||| :::::||||| :::::||||| :::::||||| :::::|||||
194 TTTACATCCAGACT.....CCGGAATTCAGTGG 222
   ||| :::::||||| :::::||||| :::::||||| :::::|||||
78 yValIleSer 81
   ::::: |||
223 CCTACCAAGT 232

```

seq\_name: gb\_gss:BH313108

seq\_documentation\_block:

LOCUS BH313108 548 bp DNA linear GSS 03-DEC-2001  
 DEFINITION CH230-101K19\_TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 CH230-101K19, DNA sequence.

ACCESSION BH313108

VERSION BH313108.1 GI:17241916

KEYWORDS GSS.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 548)  
 Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn  
 A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
 Jong,P. and Fraser,C.M

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other\_GSS: CH230-101K19.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@email.cho.org).

Clones may be purchased from BACPAC Resources

([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end

page: [http://www.tigr.org/tldb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html)

Plate: 101 row: K column: 19

Seq primer: T7

Class: BAC ends.

FEATURES Location/Qualifiers

```

1..548
   /organism="Rattus norvegicus"
   /strain="BN/SSNHsd/MCW"
   /db_xref="taxon:10116"
   /clone="CH230-101K19"
   /clone_lib="CHORI-230 Segment 1"
   /sex="Female"
   /cell_type="Brain"
   /note="Vector: PTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

```

BASE COUNT 124 a 102 c 107 g 215 t

ORIGIN

```

alignment_scores:
  Quality: 73.50 Length: 40
  Ratio: 2.625 Gaps: 2
  Percent Similarity: 70.000 Percent Identity: 45.000
alignment_block:
  US-09-528-682-3 x BH313108/rev ..

```

Align seg 1/1 to reverse of: BH313108 from: 1 to: 548

```

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79
   ||||| ||| ||||| ||||| ::::: |||||
192 ATGACACCCCTTCACATAGTCCCAATAACATAAATATCTTGGTGTGAC 143
   ||||| ||| ||||| ||||| ::::: |||||
80 .IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
142 TTTACCAAGCAAGTGAAGATCTGTATGACAAAGAACTTCAAGTCTCTGA 93
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
94 euArgPheMetGluAspAsp 100
   ::::: ||||| ||||| ||||| ||||| ||||| |||||
92 CGAAAGAAATTTGAGGAAGAT 73

```

seq\_name: gb\_est1:AUI18471

seq\_documentation\_block:

LOCUS AUI18471 852 bp mRNA linear EST 19-OCT-2000  
 DEFINITION AUI18471 HEMBAI Homo sapiens cDNA clone HEMBAI003666 5', mRNA  
 sequence.

ACCESSION AUI18471

VERSION AUI18471.1 GI:10933523

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 852)

AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
 Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
 Isogai,T.

TITLE HRI human cDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3951

Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

FEATURES Location/Qualifiers

```

1..852
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="HEMBAI003666"
   /clone_lib="HEMBAI"
   /tissue_type="whole embryo, mainly head"
   /dev_stage="embryo, 10 weeks"
   /note="Vector: pME18SFL3"

```

BASE COUNT 198 a 194 c 248 g 210 t 2 others

ORIGIN

alignment\_scores:

```

  Quality: 73.50 Length: 74
  Ratio: 1.531 Gaps: 2
  Percent Similarity: 64.865 Percent Identity: 32.432
alignment_block:

```



US-09-528-682-3 x AUL18471

Align seg 1/1 to: AUL18471 from: 1 to: 852

```

14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
311 GTTGAAGTGGCCCTACTACATTTAAAGAAGAGAGGCTTAGGCTGTG 360
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
361 GGAACCTAGTGGCCCTCCCTCCGAGGAGGAGATCTGTCAGGCAATCTG 410
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
47 lyMetGlnAsp.....GlyAspLeuPheGlyThrValAsnGlnSerAsn 61
   :::: :::: :::: :::: :::: :::: ::::
411 GGATGGTGGATGTCAGTGGACAGTTTTCGACACAGTCGATTCAGTGG 460
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
62 PheProMetSerThrPheGluGlnValProAsnAsnLysGluPheLysG 78
   :::: :::: :::: :::: :::: :::: ::::
461 TTTACATCCCACT.....CCGGAATTCAGTGG 489
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
78 yValIleSerAlaAsnVallys 85
   :::: :::: :::: :::: :::: :::: ::::
490 CCTACCAAGTGGCACAGCTGAG 511
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::

```

seq\_name: gb\_est2:BI316465

seq\_documentation\_block:

LOCUS BI316465 504 bp mRNA linear EST 29-NOV-2001  
 DEFINITION saf03f02.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 ID: Gm-cl065-4468 5' similar to TR:Q9ZW71 Q9ZW71 F14B2.18 PROTEIN.  
 ;, mRNA sequence.

ACCESSION

VERSION BI316465.1 GI:14990792

KEYWORDS EST

SOURCE soybean.

ORGANISM

Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE

AUTHORS

1 (bases 1 to 504)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna  
 , A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
 , Y., Person, B., Swallier, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
 , R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
 , R., Waterston, R. and Willson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from gibco

High quality sequence stop: 416.

FEATURES

source

1. .504

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-4468"

/clone\_lib="Gm-cl065"

/tissue\_type="germinating shoots"

/lab\_host="PH108"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI: The cDNA library was constructed from mRNA isolated

germinating shoots of the cultivar Williams. The seeds

were allowed to germinate for 24 hours prior to being

cold stressed for 2 days at 4C. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site. EcoRI  
 adapters were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the pBluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This library  
 was constructed in the laboratory of Dr. Randy  
 Shoemaker."

BASE COUNT 114 a 129 c 97 g 164 t

ORIGIN

alignment\_scores:

Quality: 73.00 Length: 99

Ratio: 1.304 Gaps: 6

Percent Similarity: 56.566 Percent Identity: 28.283

alignment\_block:

US-09-528-682-3 x BI316465/rev ..

Align seg 1/1 to reverse of: BI316465 from: 1 to: 504

```

14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
462 GTGAAGCCCACTAGAACATCAATCACCAAGCACCAAAAGACGCTTGAGTG 413
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
   :::: :::: :::: :::: :::: :::: ::::
412 CAGCTTGCTTGACGATATTCTACTTTTGTGGCGATCATTAAGCCCTTGAT 363
   :::: :::: :::: :::: :::: :::: ::::
43 ..SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
   :::: :::: :::: :::: :::: :::: ::::
362 ATAGTTTGGCGAGGG.....GGAACCAATTGGATGGCAAC 322
   :::: :::: :::: :::: :::: :::: ::::
59 GlnSer...AsnPhePro.....LysGluPheLysGlyVali 80
   :::: :::: :::: :::: :::: :::: ::::
321 ACTGCATTCATTATCCCGTCTCTTCATGTATATATCCCTCCACACTATCAGG 272
   :::: :::: :::: :::: :::: :::: ::::
66 rPheGluGlnValProAsnAsn.....LysGluPheLysGlyVali 80
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
271 GTATGAAGAACTCGATATAAATCGAGTCCAGAAGAGCATTCGGGT.... 226
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
80 leSerAlaAsnVallyTyrAspMetAsnPheLysLysLeuLeuArg 95
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::
225 .....GAAACTCGATGTGATCTGCAGCTGAGCGGCGGATTCGA 187
   :::::::::::::: :::: :::: :::: :::: :::: :::: ::::

```

seq\_name: gb\_est2:BE590377

seq\_documentation\_block:

LOCUS BE590377 565 bp mRNA linear EST 17-AUG-2000  
 DEFINITION SB117 Sugar Beet germination cDNA library Beta vulgaris cDNA clone  
 YS016232 5', mRNA sequence.

ACCESSION BE590377

VERSION BE590377.1 GI:9843416

KEYWORDS EST

SOURCE Beta vulgaris.

ORGANISM Beta vulgaris

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.

REFERENCE 1 (bases 1 to 565)

AUTHORS de los Reyes, B.G., McGrath, J.M., Myers, S. and Derrico, C.

TITLE Differential gene expression in sugar beet seedlings (Beta vulgaris

) germinated under stress conditions

JOURNAL Unpublished (2000)

COMMENT Contact: J. Mitchell McGrath

Sugar Beet Genetics Lab., USDA/ARS/Sugar Beet and Bean Research

Unit

Michigan State University

494 Plant and Soil Sci. Bldg., East Lansing MI 48824, USA

Tel: (517)-432-2355















CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
Pieter de Jong

BASE COUNT 64 a 83 c 91 g 163 t  
ORIGIN

alignment\_scores:  
Quality: 72.50 Length: 93  
Ratio: 1.394 Gaps: 3  
Percent Similarity: 55.914 Percent Identity: 26.882

alignment\_block:  
US-09-528-682-3 x BH319283/rev ..

Align seg 1/1 to reverse of: BH319283 from: 1 to: 401

```

9 GlnAlaTyrGluProIleGluValAsnThrAsnThrValThrGlnIleAs 25
|||||.....:||||:||||:||||:||||:||||:||||:||||:
277 CAGGCTTATAGGACAGTTGAACACACTGTGAGAAATAGCAACAAAGGTAA 228
25 nGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAlaLeuI 42
|||||.....:||||:||||:||||:||||:||||:||||:
227 CACCAGACAGCGCTGCTGGAGGCAAGCGCAGGAAT..... 188
42 leSerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
|||||.....:||||:||||:||||:||||:||||:||||:
187 .....GCAAGCAACAGAAACCAAGAC.....TACATGGCATCTCG 152
59 GlnSerAsnPhePrometSerThrPheGluGlnValProAsnAsnLysG 75
|||||.....:||||:||||:||||:||||:||||:||||:
151 GACCCCAATCTCCACCAAGCAACCAATGCAATATCCAAACACAAAGA 102
75 uPheLysGlyValIleSerAlaAsnValLysTyrAspMetAsnPheLys 92
|||||.....:||||:||||:||||:||||:||||:||||:
101 A.....AACAAGATCTTGATTTAAAT 79
92 ysLeuLeuArgPheMetGluAspPhe 101
::: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
78 CACATTGATCATGTATGAGGACTTC 50

```

seq\_name: gb\_est2:BE581518

seq\_documentation\_block:  
LOCUS BE581518 505 bp mRNA linear EST 09-MAY-2001  
DEFINITION K47a07.y1 TBN95TM-SSR Strongyloides stercoralis cDNA 5', mRNA  
sequence.  
ACCESSION BE581518  
VERSION BE581518.1 GI:9832460  
KEYWORDS EST.  
SOURCE Strongyloides stercoralis.  
ORGANISM Strongyloides stercoralis  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Panagrolaimoidea; Strongyloidea; Strongyloides.  
1 (bases 1 to 505)

REFERENCE  
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,  
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,  
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisshvili,R.,  
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,  
M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,  
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and  
Wilson,R.

TITLE The Washington Univ. Nematode EST Project, 1999  
JOURNAL Unpublished (1999)  
COMMENT Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

The library was constructed by Dr. Thomas Nutman and colleagues of  
NIH (tnutman@nih.gov). DNA Sequencing by: Washington  
University Genome Sequencing Center St. Louis.

High quality sequence stop: 401.  
Location/Qualifiers

```

1..505
/organism="Strongyloides stercoralis"
/strain="Rhabditiform larvae obtained from gerbils"
/db_xref="taxon:6248"
/clone_lib="XL-1 Blue MRF"
/lab_host="TBN95TM-SSR"
/Note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:  
EcoRI; Site_2: XhoI; mRNA was purified from 2 x 10B3  
rhabditiform larvae which had been isolated from gerbils  
experimentally infected with larvae originally isolated  
from experimentally infected dogs. cDNA was constructed  
and, using adaptors, was cloned unidirectionally into the  
vector from the EcoRI site to the XhoI site. The library  
has an unamplified titer of 1 x 10E5 pfu/ml and an  
amplified, undiluted titer of 9 x 10E11 pfu/ml. The  
average insert size of the unamplified library is 675 bp  
(range, 100-1700)."
```

BASE COUNT 200 a 84 c 70 g 151 t  
ORIGIN

alignment\_scores:

Quality: 72.50 Length: 97  
Ratio: 1.450 Gaps: 3  
Percent Similarity: 51.546 Percent Identity: 26.804

alignment\_block:

US-09-528-682-3 x BE581518 ..

Align seg 1/1 to: BE581518 from: 1 to: 505

```

8 GlnGluAlaTyr.....GluProIleGluValAsnThrAsnThrVa 21
||||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
33 CACACACCTTATTTGAAATTCCTCCATCACTCAATCAACATCATCATCA 82
21 lThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgTyrSerA 38
::: ||| :||| :||| :||| :||| :||| :||| :||| :|||
83 TATAAATAATAATAATAATAATAATTACCTTCAACAACAACAATCAAGTA 132
38 snPheAlaLeuIleSerAla..... 44
||||| :||| :||| :||| :||| :||| :||| :||| :|||
133 ATTAGGAATAATATGGAGCTGGAATTTGGTCTGCTAGTGGAAATTTGTT 182
45 GluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAs 61
||||| ||| :||| :||| :||| :||| :||| :||| :|||
183 CTGGAGGAAATGCTATTGGAGAT.....GGTATATCAATATATACAAA 226
61 nPhePrometSerThrPheGluGlnValProAsnAsnLysGluPheLys 78
||||| :||| :||| :||| :||| :||| :||| :||| :|||
227 TTTTGGCACAATACTACTCTATTATTAATCCTTCAACATCACCATTAGTG 276
78 lyValIleSerAlaAsnValLysTyrAspMetAsnPheLys 91
::: :||| :||| :||| :||| :||| :||| :||| :||| :|||
277 ATGCTGATACAGCTCCCAATACACATGATGATGATTATAAG 317

```

seq\_name: gb\_gss:AZ014621

seq\_documentation\_block:

LOCUS AZ014621  
DEFINITION RPCI-23-366D18.TJ RPCI-23 Mus musculus genomic clone RPCI-23-366D18  
DNA sequence.

ACCESSION AZ014621

VERSION AZ014621.1 GI:7090005

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 618)  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret  
B., Levine,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.



and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-23  
 Unpublished (1999)  
 Other GSSs: RPCI-23-366D18.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

TITLE  
 JOURNAL  
 COMMENT

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 366 row: D column: 18  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

source

Location/Qualifiers

1..618  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-366D18"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT  
 ORIGIN

247 a 112 c 107 g 152 t

## alignment\_scores:

Quality: 72.50 Length: 40  
 Ratio: 2.589 Gaps: 2  
 Percent Similarity: 70.000 Percent Identity: 45.000

## alignment\_block:

US-09-528-682-3 x AZ014621 ..

Align seg 1/1 to: AZ014621 from: 1 to: 618

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79  
 |||::: ||| ::||| ||||| ::|||: |||||  
 95 ATGACACCCCTTCACAAAGTCACAAATATACAAATATACCTTGGTGTGAC 144  
 80 .IleSerAlaAsnVallys.....TyrAspMetAsnPheLysLysLeuL 94  
 ::::: ::::: ||| ||||| ||||| |||||  
 145 TCTAACCAACAATTTGAAATCTGTATGACAAAGAACTTCAAGTCTCCTT 194  
 94 euArgPheMetGluAspAsp 100  
 ||||| ::|||: |||||  
 195 TGAGAGATTTTGAAGAAGAT 214

seq\_name: gb\_gss:BH278768

## seq\_documentation\_block:

LOCUS BH278768 637 bp DNA linear GSS 30-NOV-2001  
 DEFINITION CH230-108N7.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 CH230-108N7, DNA sequence.  
 ACCSSION BH278768  
 VERSION BH278768.1 GI:17191170  
 KEYWORDS GSS.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus

## REFERENCE

AUTHORS

1 (bases 1 to 637)

Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn  
 ,A., Gebreyeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
 Jong,P. and Fraser,C.M.  
 Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
 JOURNAL  
 COMMENT

TITLE

JOURNAL

COMMENT

Other GSSs: CH230-108N7.TJ  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/or ering\_information.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 108 row: N column: 7  
 Seq primer: T7  
 Class: BAC ends.

## FEATURES

source

Location/Qualifiers

1..637  
 /organism="Rattus norvegicus"  
 /strain="BN/SSNHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-108N7"  
 /clone\_lib="CHORI-230 Segment 1"  
 /sex="Female"  
 /cell\_type="Brain"  
 /note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
 Pieter de Jong"

BASE COUNT 252 a 122 c 112 g 151 t  
 ORIGIN

## alignment\_scores:

Quality: 72.50 Length: 40  
 Ratio: 2.589 Gaps: 2  
 Percent Similarity: 70.000 Percent Identity: 45.000

## alignment\_block:

US-09-528-682-3 x BH278768 ..

Align seg 1/1 to: BH278768 from: 1 to: 637

64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79  
 |||::: ||| ::||| ||||| ::|||: |||||  
 379 ATGACACCCCTTCACAAATAGTCCCAATATATAAATATCTTGGTGTGAC 428  
 80 .IleSerAlaAsnVallys.....TyrAspMetAsnPheLysLysLeuL 94  
 ::::: ::::: ||| ||||| ||||| |||||  
 429 TTTAACCAACAAGTGAAGATCTGTATGACAAAGAACTTCAAGTCTCTGA 478  
 94 euArgPheMetGluAspAsp 100  
 ::::: ::|||: |||||  
 479 AGAAAGAAATTTGAAGAAGAT 498

seq\_name: gb\_est2:BG524649

## seq\_documentation\_block:

LOCUS BG524649 643 bp mRNA linear EST 16-NOV-2001  
 DEFINITION 5-44 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA  
 sequence.  
 ACCSSION BG524649  
 VERSION BG524649.1 GI:16948091  
 KEYWORDS EST.



SOURCE  
ORGANISM  
Stevia rebaudiana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
Heliantheae; Stevia.  
1 (bases 1 to 643)  
Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.  
Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in  
diterpene synthesis  
Unpublished (2001)  
Contact: Jim Brandle  
Genomics and Biotechnology  
Agriculture and Agri-Food Canada - SCPFRC  
1391 Sandford St., London, Ontario, CANADA, N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3997  
Email: brandleje@agr.ca  
Seq primer: T3 promoter primer.

FEATURES  
source  
1..643  
/organism="Stevia rebaudiana"  
/strain="751/1501"  
/cultivar="Landrace"  
/db\_xref="taxon:55670"  
/clone\_lib="Stevia field grown leaf cDNA"  
/tissue\_type="leaf"  
/dev\_stage="field grown, mid-size"  
/lab\_host="E. coli strain XLOLR"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI; This  
cDNA library was constructed from polyA+ enriched mRNA  
from field grown leaves. Mid-size actively growing leaves  
were collected and pooled from several plants and frozen  
immediately after harvesting in liquid nitrogen. The cDNA  
was prepared using an xhoi-poly(dT) linker-primer. An  
EcoRI adapter was ligated to the blunt end cDNA and the  
products were digested with EcoRI and XhoI enabling  
directional cloning into the lambda ZAP Express vector.  
The library was amplified using the host strain XLI-Blue  
MRP'. Mass excision of the library was performed to  
obtain pBK-CMV phagemid clones in the host strain XLOLR.  
Single pass DNA sequencing was performed using the T3  
promoter primer: 5' ATTACCTCCTCAAGGGA 3'. This library  
was constructed by Alex Richman."

BASE COUNT 212 a 112 c 149 g 170 t  
ORIGIN  
alignment\_scores:  
Quality: 72.50 Length: 90  
Ratio: 1.450 Gaps: 3  
Percent Similarity: 55.556 Percent Identity: 25.556  
alignment\_block:  
US-09-528-682-3 x BG524649 ..  
Align seg 1/1 to: BG524649 from: 1 to: 643

24 IleAsnGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAl 40  
103 GTTAATCTGAAACATCATCGCTATTGCTGTGCTGAATCGGC 152  
40 AleuIleSerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrV 57  
153 CCATGTGCTATTGAAACGGGTCCAAAGAACGACGTTAACGGTCATG 202  
57 AlaAsnGln.....SerAsnPhe 62  
203 TTGGTGAATAATGAATCAGTTTCATGTAGAAAAAGCACACGACCATGTTT 252  
63 ProMetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVa 79  
253 CCAAAGGATGCTGTGACGAGTGGCCCTGAAGCCAAAAGTTTCATTCGGT 302

79 IleSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgp 96  
303 TTATTTT.....GTAAATATC.....CGAA 322  
96 heMetGluAspPheHe 102  
323 CAGTTGAAGACCAAGATTCGT 342  
seq\_name: gb\_gss:AZ735738

seq\_documentation\_block:  
LOCUS AZ735738 751 bp DNA linear GSS 25-JAN-2001  
DEFINITION RPCI-24-70J5.TJB RPCI-24 Mus musculus genomic clone RPCI-24-70J5,  
DNA sequence.  
ACCESSION AZ735738  
VERSION AZ735738.1 GI:12502115  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 751)  
AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,  
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,  
Russell, D., de Jong, P. and Fraser, C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-24  
JOURNAL Unpublished (1999)  
COMMENT Other\_GSSs: RPCI-24-70J5.TVB  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@igirv.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end  
page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 70 row: J column: 5  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
Location/Qualifiers  
1..751  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-70J5"  
/clone\_lib="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pRABAC1; Site\_1: BamHI; Site\_2: BamHI;  
RPCI-24 Mouse BAC Library produced by Pieter de Jong.  
The library was cloned in the pRABAC1 cloning vector at the  
BamHI sites using MboI partially digested male C57BL/6J  
DNA."

BASE COUNT 311 a 128 c 125 g 187 t  
ORIGIN

alignment\_scores:  
Quality: 72.50 Length: 73  
Ratio: 1.576 Gaps: 4  
Percent Similarity: 63.014 Percent Identity: 34.247

alignment\_block:  
US-09-528-682-3 x AZ735738 ..

Align seg 1/1 to: AZ735738 from: 1 to: 751

37 SerAsnPheAlaLeuIleSerAlaGlu.....GlyGlyMetGlnAspG1 51



```

171 AATAACTTTATTGTAGTAGCTAGATGAAATTAACCAACCAATCAGT 220
      :::::  :::::  :::::  :::::  :::::  :::::
51 YAspLeuPheGlyThrValAsnGlnSerAsnPheProMet..... 64
      |||||  |||  |||||  :::::  :::::  :::::
221 GGCCCTTTCTACACAAGAATAATCAGACTGAGAAATAATTAGTGAAA 270
      :::::  :::::  :::::  :::::  :::::  :::::
65 ..SerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal... 79
      :::::  :::::  :::::  :::::  :::::  :::::
271 CAACAACCTTCATGAGAGTCACAAATAATATAAATACTTTGGTGACT 320
      :::::  :::::  :::::  :::::  :::::  :::::
80 ILSerAlaAsnValLys.....TyrAspMetAsnPhelLysLysLeuLe 94
      :::::  :::::  :::::  :::::  :::::  :::::
321 CTATCTAAGGAAGTAAAGATCTGATGATAAGAACTCAAGTCACTGAA 370
      :::::  :::::  :::::  :::::  :::::  :::::
94 uArgPheMetGluAsp 100
      :::::  :::::  :::::  :::::  :::::  :::::
371 GAAGAATAATTGAAGAAGAT 389
      :::::  :::::  :::::  :::::  :::::  :::::

```

seq\_name: gb\_gss:AZ678077

seq\_documentation\_block:  
LOCUS AZ678077 825 bp DNA linear GSS 14-DEC-2000  
DEFINITION ENTHK25TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, DNA sequence.

ACCESSION AZ678077

VERSION AZ678077.1 GI:11815118

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 825)

AUTHORS Loftus.B., Van Aken.S. and Fraser.C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMI:IMSS sheared DNA library

COMMENT Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 17

High quality sequence stop: 823.

Location/Qualifiers

1..825

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db\_xref="taxon:5759"

/note="Vector: pHOS1; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

344 a 120 c 101 g 260 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 72.50 Length: 81  
Ratio: 1.510 Gaps: 3  
Percent Similarity: 59.259 Percent Identity: 27.160

alignment\_block:

US-09-528-682-3 x AZ678077 ..

Align seg 1/1 to: AZ678077 from: 1 to: 825

13 ProLeuGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnG1 29

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

507 CCAACTGAAGTGAATGATTAAACATGTAATGCTCTATTTCACAATTACTAA 556

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

29 uValProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluG 46

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

557 AATTCACACTAGATTTAACA..... 575

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

46 lyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPhe 62

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

576 .....ATAGAAGATGGAATGTTATTGGTACAATTACTCAATCTAGT... 617

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

63 ProMetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVa 79

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

618 GATGTTTCAACA.....CCATCAATTAATGTTGAAAAGAGGTAC 655

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

79 lIleSerAlaAsnValLysTyrAspMetAsnPhelLysLysLeu 93

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

656 ATTCACGATTAACTAAACAAGACATAACTCTTACATCAATT 698

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

seq\_name: gb\_gss:AZ536242

seq\_documentation\_block:

LOCUS AZ536242 835 bp DNA linear GSS 03-NOV-2000

DEFINITION ENT32TF Entamoeba histolytica Sheared DNA Entamoeba histolytica

genomic, DNA sequence.

ACCESSION AZ536242

VERSION AZ536242.1 GI:11093189

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 835)

AUTHORS Loftus.B., Van Aken.S. and Fraser.C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMI:IMSS sheared DNA library

COMMENT Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 27

High quality sequence stop: 803.

Location/Qualifiers

1..835

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db\_xref="taxon:5759"

/note="Vector: pHOS1; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."











## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs.r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LAM8423 row: e column: 21  
 High quality sequence stop: 682.  
 Location/Qualifiers  
 1..697  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3448316"  
 /clone\_lib="NIH\_MGC\_l0"  
 /cell\_line="MGC36"  
 /lab\_host="DH10B"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site.1: NotI;  
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.5 kb. Library prepared by Life  
 Technologies."  
 BASE COUNT 148 a 208 c 182 g 159 t  
 ORIGIN

## FEATURES

## source

FEATURES  
 source

Location/Qualifiers

1..751  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-24-287C15"  
 /clone\_lib="RPCI-24"  
 /sex="Male"  
 /cell\_type="Spleen/Brain"  
 /note="Vector: pTARBAC1; Site.1: BamHI; Site.2: BamHI;  
 RPCI-24 Mouse BAC Library produced by Pieter de Jong. The  
 library was cloned in the pTARBAC1 cloning vector at the  
 BamHI sites using MboI partially digested male C57BL/6J  
 DNA."  
 BASE COUNT 175 a 126 c 150 g 300 t  
 ORIGIN

## alignment\_scores:

Quality: 71.50 Length: 92  
 Ratio: 1.375 Gaps: 4  
 Percent Similarity: 56.522 Percent Identity: 31.522

## alignment\_block:

US-09-528-682-3 x BH096764/rev ..

Align seg 1/1 to reverse of: BH096764 from: 1 to: 751

13 ProlleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnG1 29  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 705 CCAATATTCACAGAGAACTCTAAACCTGTATAACAGTTCATGA 656  
 29 uValProLeuAspGlyArgTyrSerAsn.....PheAlaLeuIleSerA 44  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 655 AGTAGCTGGATATAAAATTAACCTCAAAAGTCAATGGCCTTCTCTACA 606  
 44 laGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSer 60  
 : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 605 CAAAGGATAACAGCGCTGAGAAGAAATTAGGGGAACA..... 568  
 61 AsnPheProMetSerThrPheGluGlnValProAsnAsnLysGluPheLy 77  
 : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 567 .....ACACCTTCTCAATAGTCACAAATAATAATAATACCT 530  
 77 sGlyVal...IleSerAlaAsnValLys.....TyrAspMetAsnPhel 91  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 529 TGGCGTGACTCTAACTAAGGAAGTGAAGATCTGTATGATAAGAAATTTCA 480  
 91 yLysLeuLeuArgPheMetGluAsp 99  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 479 AGTCTCTGAAGAAGATATTAAAGAA 454

seq\_name: gb\_gss:CNS05420

seq\_documentation\_block:

LOCUS CNS05420 1005 bp DNA linear GSS 26-JUL-2000  
 DEFINITION Tetraodon nigroviridis genome survey sequence f3 end of clone  
 053D17 of library A from Tetraodon nigroviridis, genomic survey

seq\_name: gb\_gss:BH096764

seq\_documentation\_block:

LOCUS BH096764 751 bp DNA linear GSS 19-JUL-2001  
 DEFINITION RPCI-24-287C15.TV RPCI-24 Mus musculus genomic clone RPCI-24-287C15  
 , DNA sequence.

## ACCESSION

BH096764

## VERSION

BH096764.1 GI:14918110

## KEYWORDS

GSS.

## SOURCE

house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 751)  
 Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,  
 Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,  
 Russell,D., de Jong,P. and Fraser,C.M.  
 Mouse BAC End Sequences from Library RPCI-24  
 Unpublished (1999)  
 Other\_GSSs: RPCI-24-287C15.TJ



```

sequence.
ACCESSION      AL320217
VERSION        AL320217.1 GI:9553101
KEYWORDS       GSS: genome survey sequence.
SOURCE         Tetraodon nigroviridis.
ORGANISM       Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 1005)
AUTHORS       Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
              Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
              Saurin,W. and Weissenbach,J.
TITLE         Estimate of human gene number provided by genome-wide analysis
              using Tetraodon nigroviridis DNA sequence
JOURNAL        Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE        20996633
REFERENCE      2 (bases 1 to 1005)
AUTHORS       Crolius,H.R., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C.,
              Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W.,
              Bernot,A. and Weissenbach,J.
TITLE         Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish tetraodon nigroviridis
JOURNAL        Genome Res. 10 (7), 939-949 (2000)
MEDLINE        20359837
REFERENCE      3 (bases 1 to 1005)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL        Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT        This sequence is a single read and was generated as part of a large
              scale clone-and-sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/Tetraodon.
FEATURES       location/Qualifiers
              1..1005
               /organism="Tetraodon nigroviridis"
               /db_xref="taxon:99883"
               /clone="053D17"
               /clone_lib="A"
               /note="Genoscope sequence ID : C0AA053CB09A1-end : T3"
BASE COUNT     210 a 252 c 268 g 263 t 12 others
ORIGIN
alignment_scores:
  Quality: 71.50      Length: 72
  Ratio: 1.625        Gaps: 1
  Percent Similarity: 61.111 Percent Identity: 29.167
alignment_block:
US-09-528-682-3 x CNS05420/rev ..
Align seg 1/1 to reverse of: CNS05420 from: 1 to: 1005
14  lIeGlUValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 ATCATGATGAATCACTACTACTACTACTATGCGCAATGACCTCTAGTTC 202
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30  lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 ACCA.....AACCTCTATCGCTCACCCGCCACTCAAGTGGG 167
47  lYMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPhePro 63
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 GATTGTTCCTGGGGATCTGTTGGAGGAGGGAATGGGGGTCCACTGGCCA 117
64  MetSerThrPheGlnGlnValProAsnAsnLysGluPheLysGlyValIi 80
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 ATGGAACCTTTCAACCCCTCAGCAACCCCAAGGAGAACTGGGGGGTCT 67
80  eSerAlaAsnValLys 85
|||  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

seq_name: gb_gss:A2329751
seq_documentation_block:
LOCUS      A2329751
DEFINITION 1M0054F14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0054F14 R, DNA sequence.
ACCESSION  A2329751
VERSION    A2329751.1 GI:10390776
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 485)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Bescorn,T., Duval,B., Hamill,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
          ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Insert Length: 10000 Std Error: 0.00
          Plate: 0054 row: F column: 14
          Seq primer: CACACAGGAACAGCTATGACC
          Class: plasmid ends
          High quality sequence stop: 485.
FEATURES   Location/Qualifiers
              1..485
               /organism="Mus musculus"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="UUGC1M0054F14"
               /clone_lib="Mouse 10kb plasmid UUGC1M library"
               /sex="Male"
               /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
               /note="Vector: PWB42nv; Purified genomic DNA from M.
               musculus C57BL/6J (male) was obtained from the Jackson
               Laboratory Mouse DNA Resource
               (http://www.jax.org/resources/documents/dnares/). The DNA
               was hydrodynamically sheared by repeated passage through a
               0.005 inch orifice at constant velocity. The sheared DNA
               was blunt end-repaired with T4 DNA polymerase and T4
               polynucleotide kinase. Adaptor oligonucleotides were
               ligated to the blunt ends in high molar excess. The
               adaptor DNA was purified and size-selected for a 9.5 to
               10.5 kb range using preparative agarose gel
               electrophoresis. Vector DNA was prepared from a derivative
               of PWB42 (GI14732114)gB/AFL29072.1), a copy-number
               inducible derivative of plasmid R1. The vector was ligated
               with adaptors complementary to the insert adaptors and
               purified. The sheared, adaptor mouse DNA was annealed to
               adaptor vector DNA, and transformed into
               chemically-competent E. coli XL10-Gold (Stratagene) cells
               and selected for ampicillin resistance."
BASE COUNT 113 a 83 c 94 g 195 t
ORIGIN
alignment_scores:
  Quality: 71.00      Length: 62
  Ratio: 1.919        Gaps: 3
  Percent Similarity: 59.677 Percent Identity: 32.258

```



## alignment\_block:

US-09-528-682-3 x AZ329751/rev ..

Align seg 1/1 to reverse of: AZ329751 from: 1 to: 485

```

57 ValAsnGlnSerAsnPhePrometSerThr..... 66
:|||||:|||||:|||||:|||||:
465 CTCAACAAGTCAATGGCTTCTCTACACAAGATAAACAGGCTGAG 416
77 .....PheGluGlnValProAsnAsnLysG 75
|||||:|||||:|||||:|||||:
415 AAAGAAATTAGGGAACAACCCCTCTCAATAGTCCCAATAATATAA 366
75 luPheLysGlyVal...IleSerAlaAsnValLys.....TyrAspMet 88
:|||||:|||||:|||||:|||||:
365 AATACCTTGGAGTGACTCTAACTAAGCAACTGAAGATCTGTATGATAAG 316
89 AsnPhelLysLysLeuLeuArgPheMetGluAspAsp 100
|||||:|||||:|||||:|||||:
315 AACTTCAAGTCTCTGAAGAAAGAAATTAAGAAGAT 280

```

seq\_name: gb\_gss:AZ515059

## seq\_documentation\_block:

LOCUS AZ515059 485 bp DNA linear GSS 05-OCT-2000  
DEFINITION 1M0054F14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0054F14 F, DNA sequence.

ACCESSION AZ515059

VERSION AZ515059.1 GI:10696375

## KEYWORDS

SOURCE house mouse.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 485)  
Dunn,D., Aoyagi,A., Barber,M., Bescorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

## TITLE

Unpublished (2000)

## JOURNAL

## COMMENT

Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0054 row: F column: 14  
Seq primer: CGTGTAAACACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 485.

Location/Qualifiers

1. .485

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0054F14"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The

## FEATURES

## source

adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent *E. coli* XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 113 a 83 c 94 g 195 t

## ORIGIN

## alignment\_scores:

Quality: 71.00 Length: 62  
Ratio: 1.919 Gaps: 3  
Percent Similarity: 59.677 Percent Identity: 32.258

## alignment\_block:

US-09-528-682-3 x AZ515059/rev ..

Align seg 1/1 to reverse of: AZ515059 from: 1 to: 485

```

57 ValAsnGlnSerAsnPhePrometSerThr..... 66
:|||||:|||||:|||||:|||||:
465 CTCAACAAGTCAATGGCTTCTCTACACAAGATAAACAGGCTGAG 416
67 .....PheGluGlnValProAsnAsnLysG 75
|||||:|||||:|||||:|||||:
415 AAAGAAATTAGGGAACAACCCCTCTCAATAGTCCCAATAATATAA 366
75 luPheLysGlyVal...IleSerAlaAsnValLys.....TyrAspMet 88
:|||||:|||||:|||||:|||||:
365 AATACCTTGGAGTGACTCTAACTAAGCAACTGAAGATCTGTATGATAAG 316
89 AsnPhelLysLysLeuLeuArgPheMetGluAspAsp 100
|||||:|||||:|||||:|||||:
315 AACTTCAAGTCTCTGAAGAAAGAAATTAAGAAGAT 280

```

seq\_name: gb\_gss:BH600950

## seq\_documentation\_block:

LOCUS BH600950 502 bp DNA linear GSS 15-DEC-2001  
DEFINITION BOGZV06TR BOGZ Brassica oleracea genomic clone BOGZV06, DNA  
sequence.

ACCESSION BH600950

VERSION BH600950.1 GI:17853396

## KEYWORDS

SOURCE Brassica oleracea.

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 502)  
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)

Contact: Chris Town

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. .502

/organism="Brassica oleracea"

/strain="T0100DH3"

/db\_xref="taxon:3712"











Ratio: 1.392 Gaps: 3  
Percent similarity: 68.919 Percent Identity: 27.027

## alignment\_block:

US-09-528-682-3 x A2672583 ..

Align seg 1/1 to: A2672583 from: 1 to: 871

```

20 ThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgTy 36
   :::::||||| :::::||||| :::::||||| :::::|||||
375 TCATCACTCTTCCACTGGAATGAATGATTTAACTGATGCTGCTATT 424
   :::::||||| :::::||||| :::::||||| :::::|||||
36 rSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspGlyAspL 53
   :::::||||| :::::||||| :::::||||| :::::|||||
425 CACA...ATTACTAAATCCACTAGATTTAACAAATAGAAGATGAAATG 471
   :::::||||| :::::||||| :::::||||| :::::|||||
53 euPheGlyThrValAsnGlnSerAsnPheProMetSerThrPheGluGln 69
   :::::||||| :::::||||| :::::||||| :::::|||||
472 TTATTGTTACAATTACTCAATCTAGT...GATGTTTCAACA..... 509
   :::::||||| :::::||||| :::::||||| :::::|||||
70 ValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLysTy 86
   :::::||||| :::::||||| :::::||||| :::::|||||
510 ...CCATCAATTAATGTTGAAAAGGTACATTCACGATTAACTAAACA 556
   :::::||||| :::::||||| :::::||||| :::::|||||
86 rAspMetAsnPheLysLysLeu 93
   :::::||||| :::::||||| :::::||||| :::::|||||
557 AGACATAACTCTTACATCAATT 578
   :::::||||| :::::||||| :::::||||| :::::|||||

```

seq\_name: gb\_estl:AW587961

```

LOCUS      AW587961                      438 bp  mRNA  linear  EST 09-MAY-2001
DEFINITION  kp10c06.y1 TBN95TM-SSFH Strongyloides stercoralis cDNA 5', mRNA
ACCESSION  AW587961
VERSION    AW587961.1  GI:7274991
KEYWORDS   Strongyloides stercoralis
SOURCE     Strongyloides stercoralis
ORGANISM   Strongyloides stercoralis
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Panagrolaimoidea; Strongyloidea; Strongyloides.
McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,
Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,
Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R.,
Ronko, I., Kennedy, S., McGuire, L., Beck, C., Underwood, K., Steptoe
, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.
, Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Dr. Thomas Nutman and colleagues of
NIH/ID, NIH (nutman@nih.gov). DNA Sequencing by: Washington
University Genome Sequencing Center St. Louis.
Seq primer: T3 ET from Amersham
High quality sequence stop: 418.
Location/Qualifiers
1..438
/organism="Strongyloides stercoralis"
/strain="Filariform larvae obtained from humans"
/db_xref="taxon:6248"
/clone_lib="TBN95TM-SSFH"
/lab_host="xL-1 Blue MRF" (Stratagene)
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:
ECORI; Site_2: XhoI; mRNA was purified from 4 x 10E5
filariform larvae which had been isolated from infected

```

## FEATURES

source

```

1..438
/organism="Strongyloides stercoralis"
/strain="Filariform larvae obtained from humans"
/db_xref="taxon:6248"
/clone_lib="TBN95TM-SSFH"
/lab_host="xL-1 Blue MRF" (Stratagene)
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:
ECORI; Site_2: XhoI; mRNA was purified from 4 x 10E5
filariform larvae which had been isolated from infected

```

humans. cDNA was constructed and, using adaptors, was cloned unidirectionally into the vector from the EORI site to the XhoI site. The library has an unamplified titer of 1.5 x 10E6 pfu/ml and an amplified, undiluted titer of 7 x 10E9 pfu/ml. The average insert size of the unamplified library is 975 bp (range, 500-1500)."

BASE COUNT 211 a 37 c 53 g 137 t  
ORIGIN

## alignment\_scores:

Quality: 70.50 Length: 92  
Ratio: 1.195 Gaps: 3  
Percent Similarity: 64.130 Percent Identity: 25.000

## alignment\_block:

US-09-528-682-3 x AW587961 ..

Align seg 1/1 to: AW587961 from: 1 to: 438

```

17 AsnThrAsnThrValThrGlnIleAsnGlySerAsnGlu.....ValPr 31
   |||:::||||| :::::||||| :::::||||| :::::|||||
77 AATAATAATACACGAAAGATAAAATTTGAGAACACAGAATATATCA 126
   :::::||||| :::::||||| :::::||||| :::::|||||
31 oLeuAspGlyArgTy rSerAsnPheAlaLeuIle.....SerA 44
   :::::||||| :::::||||| :::::||||| :::::|||||
127 TTTCCCGCAGCGCATGTCAANTACTTCAATAATAAAGATTGGATGATA 176
   :::::||||| :::::||||| :::::||||| :::::|||||
44 laGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSer 60
   :::::||||| :::::||||| :::::||||| :::::|||||
177 AAAAGAGAAAAATCGGTGAAAAACAACTTTTAAATAAGTAAAAAATAAA 226
   :::::||||| :::::||||| :::::||||| :::::|||||
61 AsnPheProMetSerThrPheGluGlnValProAsnAsnLysGluPheLy 77
   :::::||||| :::::||||| :::::||||| :::::|||||
227 AGTTTAAAAATAATAATTTTAAAAAATGTTAATAATAATAAAGAC...AA 273
   :::::||||| :::::||||| :::::||||| :::::|||||
77 sGlyValIleSerAlaAsnValLysTy rAspMetAsnPheLysLysLeuL 94
   |||:::||||| :::::||||| :::::||||| :::::|||||
274 AACTATAATCTCGAATAATATATAGGAAGAATTCCTTACTTCAAAATAGTT 323
   :::::||||| :::::||||| :::::||||| :::::|||||
94 euArgPheMetGluAspAspPheIle 102
   :::::||||| :::::||||| :::::||||| :::::|||||
324 ATAAGGTAATAAAAAATGATACCATTT 349
   :::::||||| :::::||||| :::::||||| :::::|||||

```

seq\_name: gb\_estl:AA823163

## seq\_documentation\_block:

```

LOCUS      AA823163                      448 bp  mRNA  linear  EST 17-FEB-1998
DEFINITION  vw41e03.r1 Soares_mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:1246396 5' similar to SW:Y079_HUMAN P53992 HYPOTHETICAL
PROTEIN KIAA0079 ;, mRNA sequence.

```

ACCESSION AA823163

VERSION AA823163.1 GI:2893031

## KEYWORDS

EST.

## SOURCE

house mouse.

## ORGANISM

Mus musculus.

## REFERENCE

1 (bases 1 to 448)

## AUTHORS

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu



This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:660084  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 390.

FEATURES

Location/Qualifiers  
1. .448  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1246396"  
/clone\_lib="Soares\_mammary\_gland\_NBMG"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Organ: mammary gland; Vector: pT7f3D-Pac (Pharmacia ) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5']  
TGTACCAATCTGAAGTGGGAGCGCGGAATGTTTTTTTTTTTTTTTTTTT  
T 3'); double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT7f3 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 105 a 111 c 119 g 113 t  
ORIGIN

alignment\_scores:  
Quality: 70.50 Length: 128  
Ratio: 1.102 Gaps: 3  
Percent Similarity: 50.000 Percent Identity: 21.094

alignment\_block:

US-09-528-682-3 x AA823163 ..  
Align seg 1/1 to: AA823163 from: 1 to: 448  
3 PheThrArgAlaLeuGlnGlnAlaTyGluProIleGluValAsnThrAs 19  
10 TATGTCGGCAGCTGGTTCTCCATCGATGTGGCTGAGACCAATGCTT 59  
19 nThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgT 36  
60 CTTCTACCTCGGCTCTACCACATGACAAAGTCTCCCTTGATAGTACCG 109  
36 YrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52  
110 CTGAACACCAAGAGTTCGAGCTCTGAAGAGCGCCTAAGCAGTGGCGAT 159  
53 LeuPhe.....  
160 ATATATTGCTGGAATATGGCTCAACCTCTTCTGCTGGTGGGAGCAAG 209  
55 .....GlyThrValAsnGlnSerAsnPheProMetSerThrPheG 68  
210 CGTCAACAAGGTGTGTC....CAGAGCCCTTTCACAGTGTCTCCCTTCA 256  
68 luGlnVal.....ProAsn 72  
257 GTCAGATCACCAGCGGCTTGAGTGTCTGCCAGTCTCTGGATAATCCACTG 306  
73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyAspMetAs 89  
307 TCTAAGAAGTTCGAGGTCTCATGATAGTCTACGGGCACAGAGAATGCG 356  
89 nPheLysLysLeuLeuArgPheMetGluAspAsp 100  
357 GTACATGAGCTTATGCTGGTCAAGCAGGAGAGAC 390

seq\_name: gb\_gss:BH295572

seq\_documentation\_block:

LOCUS BH295572 515 bp DNA linear GSS 30-NOV-2001  
DEFINITION CH230-187A17.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
CH230-187A17, DNA sequence.  
ACCESSION BH295572  
VERSION BH295572.1 GI:17207980.  
KEYWORDS GSS.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE

1 (bases 1 to 515)  
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn  
A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
Jong,P. and Fraser,C.M.  
Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
Unpublished (1999)  
CONTACT: Shaying Zhao  
DEPARTMENT: Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

TITLE

JOURNAL  
COMMENT  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering\_information.htm). BAC end  
plate: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
Plate: 187 row: A column: 17  
Seq primer: 17  
Class: BAC ends.

FEATURES

Location/Qualifiers  
1..515  
/organism="Rattus norvegicus"  
/strain="BN/SSNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-187A17"  
/clone\_lib="CHORI-230 Segment 1"  
/sex="Female"  
/cell\_type="Brain"  
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
Pieter de Jong"  
BASE COUNT 205 a 98 c 92 g 120 t  
ORIGIN

alignment\_scores:  
Quality: 70.50 Length: 40  
Ratio: 2.611 Gaps: 2  
Percent Similarity: 67.500 Percent Identity: 45.000

alignment\_block:

US-09-528-682-3 x BH295572 ..  
Align seg 1/1 to: BH295572 from: 1 to: 515  
64 MetSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal.. 79  
359 ATGTCACCCCTTCATAATAGTCCCAAAATATAATAACATACCTCGGTGTGAC 408  
80 .IleSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuL 94  
409 TTTTAAACAAGAGTGAAGATCTGTATGATAGATAAGAACTTCAAGCCTATGA 458  
94 euArgPheMetGluAspAsp 100  
459 AGAAGAAATTTGAAGAGAC 478



```

seq_name: gb_gss:AZ423995

seq_documentation_block:
LOCUS      AZ423995                535 bp      DNA      linear      GSS 03-OCT-2000
DEFINITION  IM0203B13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0203B13 F, DNA sequence.
ACCESSION   AZ423995
VERSION     AZ423995.1 GI:10548008
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 535)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE       Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL     Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0203 row: B column: 13
            Seq primer: CGTTGTAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 535.
FEATURES             Location/Qualifiers
     source           1..535
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone_lib="UUGC1M0203B13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="vector: pWD42nv; Purified genomic DNA from M.
            musculus C57BL/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a
            0.005 inch orifice at constant velocity. The sheared DNA
            was blunt end-repaired with T4 DNA polymerase and T4
            polynucleotide kinase. Adaptor oligonucleotides were
            ligated to the blunt ends in high molar excess. The
            adapted DNA was purified and size-selected for a 9.5 to
            10.5 kb range using preparative agarose gel
            electrophoresis. Vector DNA was prepared from a derivative
            of pWD42 (gi14732114|gb|AF129072.1), a copy-number
            inducible derivative of plasmid R1. The vector was ligated
            with adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adapted vector DNA, and transformed into
            chemically-competent E. coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
BASE COUNT      224 a 100 c 86 g 125 t
ORIGIN

alignment_scores:
Quality: 70.50      Length: 39
Ratio: 2.518       Gaps: 2
Percent Similarity: 71.795      Percent Identity: 43.590

alignment_block:
3 PheThrArgAlaLeuGlnGlnAlaTyrGluProlleGluValAsnThrAs 19
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
77 TATGTTGGCAGCTGGTTCCTCCATGGATGGCTGAGACCAATGCTCT 126
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
19 nThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArg 36
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
127 CTTCTACCTCGGCTCCTACCACTGACAAGTCTCCCTTGATAGTACCG 176
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
36 YrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
177 CTGAACCAACGACGAGTTCGAGCTCTCTGAAGAGCGCGCTTAAGCAGTGC 226
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::

```

```

US-09-528-682-3 x AZ423995 ..
Align seg 1/1 to: AZ423995 from: 1 to: 535

65 SerThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal...11 80
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
163 ACAACCTTCATATATCATCAATAATAGAAAATACCTTGGAGTGACTCT 212
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
80 eSerAlaAsnValLys.....TyrAspMetAsnPheLysLysLeuLeuA 95
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
213 AACTTAGGAGTGAAGATCTGTATGATAAAACCTTCAAGTCTCTGAAGA 262
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::

95 rgPheMetGluAspAsp 100
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
263 GAGAAATCGAAGAAGAT 279
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::

seq_name: gb_est2:BG804725

seq_documentation_block:
LOCUS      BG804725                600 bp      mRNA      linear      EST 20-DEC-2001
DEFINITION  0271-56 Mouse E14.5 retina lambda ZAP II Library Mus musculus CDNA,
            mRNA sequence.
ACCESSION   BG804725
VERSION     BG804725.1 GI:17951649
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 600)
AUTHORS     Mu,X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W.,
            White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
TITLE       Gene expression in the developing mouse retina by EST sequencing
            and microarray analysis
JOURNAL     Nucleic Acids Res. 29 (24), 4983-4993 (2001)
COMMENT     Contact: Klein WH
            Department of Biochemistry and Molecular Biology
            University of Texas M.D. Anderson Cancer Center
            Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
            Tel: 713 792 3646
            Fax: 713 790 0329.
FEATURES             Location/Qualifiers
     source           1..600
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone_lib="Mouse E14.5 retina lambda ZAP II Library"
                     /tissue_type="neural retina"
                     /dev_stage="embryonic day 14.5 post-fertilization"
BASE COUNT      142 a 143 c 161 g 154 t
ORIGIN

alignment_scores:
Quality: 70.50      Length: 128
Ratio: 1.102       Gaps: 3
Percent Similarity: 50.000      Percent Identity: 21.094

alignment_block:
US-09-528-682-3 x BG804725 ..
Align seg 1/1 to: BG804725 from: 1 to: 600

3 PheThrArgAlaLeuGlnGlnAlaTyrGluProlleGluValAsnThrAs 19
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
77 TATGTTGGCAGCTGGTTCCTCCATGGATGGCTGAGACCAATGCTCT 126
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
19 nThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArg 36
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
127 CTTCTACCTCGGCTCCTACCACTGACAAGTCTCCCTTGATAGTACCG 176
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
36 YrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
177 CTGAACCAACGACGAGTTCGAGCTCTCTGAAGAGCGCGCTTAAGCAGTGC 226
   :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::

```











```

alignment_scores:
  Quality: 70.50      Length: 128
  Ratio: 1.102       Gaps: 3
  Percent Similarity: 50.000   Percent Identity: 21.094

alignment_block:
  US-09-528-682-3 x BG865274 ..
  Align seg 1/1 to: BG865274 from: 1 to: 774

    3 PheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAs 19
    140 TATGTCGGCAGCTGGTTCTCCATCGATGTGGCTGAGACCAATGCTT 189
    19 nThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgT 36
    190 CTTCTACCTCGGCTCTACCACTGACAAAGTCTCCCTTGATAGTACCG 239
    36 yrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52
    240 CTGACACCACGAGCTGCGACCTCTGAAGAGCGCCTAAGCAGTGGCGAT 289
    53 LeuPhe..... 54
    290 ATATATTGCTGGAATGGGCTCAACCTCTTTGCTGGTGGGAGCAAG 339
    55 .....GlyThrValAsnGlnSerAsnPheProMetSerThrPheG 68
    340 CGTCCAAACAAGGTGTGTC...CAGAGCCTTTTCAACGTGTCTCCTCTCA 386
    68 luGlnVal.....ProAsn 72
    387 GTCAGATCACCAGCGGCTTGAGTGTTCGCCAGTCTGGATATCCACTG 436
    73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAspMetas 89
    437 TCTAAGAAAGTTCGAGGTCTCATCGATAGCTTAGGGGCACAGAGAAATCG 486
    89 nPheLysLysLeuLeuArgPheMetGluAspAsp 100
    487 GTACATGAAGCTTATCGTGGTCAAGCAGGAGAGAC 520

seq_name: gb_est2:BI696195

seq_documentation_block:
LOCUS      BI696195          832 bp      mRNA      linear      EST 18-SEP-2001
DEFINITION 603345846F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5373466 5',
            mRNA sequence.
ACCESSION  BI696195
VERSION    BI696195.1 GI:15658824
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 832)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLML1950 row: 1 column: 11
            High quality sequence stop: 817.
            Location/Qualifiers

FEATURES

```

```

source
1..832
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5373466"
/clonelib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT  198 a 199 c 217 g 218 t
ORIGIN

alignment_scores:
  Quality: 70.50      Length: 128
  Ratio: 1.102       Gaps: 3
  Percent Similarity: 50.000   Percent Identity: 21.094

alignment_block:
  US-09-528-682-3 x BI696195 ..
  Align seg 1/1 to: BI696195 from: 1 to: 832

    3 PheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAs 19
    63 TATGTCGGCAGCTGGTTCTCCATCGATGTGGCTGAGACCAATGCTT 112
    19 nThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgT 36
    113 CTTCTACCTCGGCTCTACCACTGACAAAGTCTCCCTTGATAGTACCG 162
    36 yrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52
    163 CTGAACACCACGAGCTGCGACCTCTGAAGAGCGCCTAAGCAGTGGCGAT 212
    53 LeuPhe..... 54
    213 ATATATTGCTGGAATGGGCTCAACCTCTTTGCTGGTGGGAGCAAG 262
    55 .....GlyThrValAsnGlnSerAsnPheProMetSerThrPheG 68
    263 CGTCCAAACAAGGTGTGTC...CAGAGCCTTTTCAACGTGTCTCCTCTCA 309
    68 luGlnVal.....ProAsn 72
    310 GTCAGATCACCAGCGGCTTGAGTGTTCGCCAGTCTGGATATCCACTG 359
    73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAspMetas 89
    360 TCTAAGAAAGTTCGAGGTCTCATCGATAGCTTAGGGGCACAGAGAAATCG 409
    89 nPheLysLysLeuLeuArgPheMetGluAspAsp 100
    410 GTACATGAAGCTTATCGTGGTCAAGCAGGAGAGAC 443

```











OM of: US-09-528-682-3 to: Issued\_Patents\_NA:\* out\_format : pfs

Date: Jun 18, 2002 7:41 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet\_p2n.model -DEV=xlip  
-Q/cgn2\_1/USPTO\_spool/US09528682/runat\_18062002\_082443\_7827/app\_query.fasta\_1.689  
-DB=Issued\_Patents\_NA -QFM=fastap -SUFFIX=p2n.rni  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -OGAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000  
-XGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALLIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFM=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682@cgn1\_1\_96 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT\_THREADS=1

Search information block:

Query: US-09-528-682-3

Query length: 107

Database: Issued\_Patents\_NA.\*

Database sequences: 383533

Database length: 122816732

Search time (sec): 143.210000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-216-276A-30 +	77.00	163.04	0.2155	1622	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-216-276A-18 +	77.00	155.70	0.5523	3230	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-219-262B-11 +	77.00	155.70	0.5523	3230	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-031-655-11 +	77.00	155.70	0.5523	3230	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-944-943-1 +	74.00	148.36	1.42	3180	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-232-479-13 +	73.00	145.60	2.02	3257	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-708-541A-27 +	73.00	145.59	2.02	3261	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-940-968-23 +	73.00	145.30	2.10	3350	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-023-610-1 +	73.00	145.30	2.10	3350	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-288-065A-1 +	73.00	145.30	2.10	3350	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-362-240A-1 +	73.00	145.30	2.10	3350	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-530-492-2 +	71.00	146.16	1.88	1931	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-906-517-2 +	70.00	132.71	10.54	5398	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-103-840A-2 -	70.00	61.31	9.4e+04	4403765	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-944-525-1 +	69.00	135.84	7.05	3180	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-219-262B-13 +	69.00	135.84	7.05	3180	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-031-655-13 +	69.00	135.84	7.05	3180	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-708-541A-31 +	69.00	135.57	7.31	3264	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-910-501-3 +	68.50	147.62	1.56	936	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-398-550-3 +	68.50	147.62	1.56	936	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-910-501-1 +	68.50	147.58	1.56	939	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-216-276A-15 +	68.00	141.55	3.39	1471	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-379-523-4 +	66.50	129.01	16.94	3356	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-936-165A-232 +	65.50	146.41	1.82	518	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-936-165A-232 +	65.50	140.48	3.89	904	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-676-967-4 -	65.00	129.39	16.14	2277	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-676-974-4 -	65.00	129.39	16.14	2277	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-098-487-4 -	63.00	108.53	234.42	10091	

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-802-141-3 +	62.50	122.79	37.65	2352	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-718-575-3 +	62.50	118.54	64.92	3505	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-269-3 +	62.50	118.54	64.92	3505	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-486-286A-3 +	62.50	118.54	64.92	3505	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-397-602A-8 +	62.00	126.67	22.98	1457	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-685-774-1 +	62.00	126.67	22.98	1457	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-282-147-36 +	61.50	117.35	75.58	3097	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-540-804-13 +	61.50	112.58	139.43	4849	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-218-265-13 +	61.50	112.58	139.43	4849	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-521-872-13 +	61.50	112.58	139.43	4849	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-590-339-13 +	61.50	112.58	139.43	4849	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-688-988-15 +	61.00	139.02	4.69	360	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-936-165A-234 +	61.00	132.66	10.61	654	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-936-165A-134 +	61.00	132.66	10.61	654	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-259-000-4 +	60.50	114.24	112.69	3280	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-729-767-6 +	60.50	114.24	112.69	3280	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-858-207A-226 +	60.00	128.96	17.05	732	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-973-462-2 +	60.00	107.75	258.85	5361	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-326-117B-1 +	60.00	107.33	273.20	5577	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-982-129-1 +	60.00	107.33	273.20	5577	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-973-462-1 +	60.00	106.29	312.39	6152	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-385-335A-9 +	60.00	105.95	326.21	6350	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-557-852-6 +	59.50	134.72	8.15	379	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-387-858A-6 +	59.50	134.72	8.15	379	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-294-384B-6 +	59.50	134.72	8.15	379	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-938A-7 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-938A-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-910-041-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-328-474-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-100-546-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-010-715-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-577-758-30 +	59.50	121.31	45.50	1335	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-470-187-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-318-905-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-232-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-140-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-938A-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-910-041-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-328-474-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-100-546-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-010-715-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-577-758-7 +	59.50	119.93	54.33	1520	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-169-927-1 +	59.50	106.59	300.68	5319	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-569-166-28 +	59.00	132.07	11.44	432	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-998-416-256 -	59.00	126.78	22.55	710	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-446-504-18 +	59.00	119.30	58.91	1434	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-712-266-18 +	59.00	119.30	58.91	1434	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-446-504-17 +	59.00	115.70	93.37	2009	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-712-266-17 +	59.00	115.70	93.37	2009	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-442-100-3 +	59.00	108.41	237.88	3984	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-618-946B-24 +	58.50	113.40	135.42	2217	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-618-946B-25 +	58.50	112.37	143.12	2442	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-400-087-2 +	58.50	111.51	159.94	2649	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-687-379-5 +	58.50	110.34	185.70	2955	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-687-379-7 +	58.50	110.34	185.70	2955	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-327-1 +	58.50	108.83	225.56	3407	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-666-221B-1 +	58.50	108.83	225.56	3407	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-666-221B-2 +	58.50	108.83	225.56	3407	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-618-946B-24 +	58.50	106.97	286.03	4054	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-793-824-1 +	58.50	105.06	365.71	4853	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-953-376-10 +	58.00	122.45	39.30	843	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-676-818-1 +	58.00	97.87	920.04	8478	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-318-905-23 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-232-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-140-23 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-938A-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-910-041-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-328-474-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-100-546-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-010-715-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-577-758-25 +	57.50	126.41	23.66	517	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-358-117-1 +	57.50	111.69	156.18	2058	















[illegible]



/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-929-329-2 +	50.00	88.47	3.1e+03	3126	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-993-825-1 -	49.50	83.87	5.5e+03	4279
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-686-968-1 +	50.00	87.34	3.1e+03	3164	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-726-214-15 -	49.50	83.10	6.1e+03	4601
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-637-732A-1 +	50.00	88.99	3.3e+03	3270	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-038-768A-1 +	49.50	82.14	6.9e+03	5033
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-100-709-3 +	50.00	86.02	4.2e+03	3934	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-851-843A-68 -	49.50	81.11	7.8e+03	5544
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-176-865-3 +	50.00	86.02	4.2e+03	3934	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-974-349A-111 -	49.50	81.11	7.8e+03	5544
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-474-038-3 +	50.00	86.02	4.2e+03	3934	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-854-050-68 -	49.50	81.11	7.8e+03	5544
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-779-046-3 +	50.00	86.02	4.2e+03	3934	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-430-323-68 -	49.50	81.11	7.8e+03	5544
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-881-340-3 +	50.00	86.02	4.2e+03	3934	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-233-086-2 -	49.50	80.57	8.4e+03	5836
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-532-547-4 +	50.00	84.96	4.8e+03	4344	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-11684-3 +	49.50	80.18	8.8e+03	60
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-379-656B-4 +	50.00	84.96	4.8e+03	4344	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-318-448-11 -	49.50	68.22	4.0e+04	18596
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-455-838-4 +	50.00	84.96	4.8e+03	4344	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-844-056-3 -	49.00	109.47	207.75	344
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-019-809-4 +	50.00	84.96	4.8e+03	4344	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-385 +	49.00	109.35	211.05	346
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-308-872B-3 +	50.00	84.76	4.9e+03	4429	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-171-945-23 +	49.00	108.38	238.86	381
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-207-999-1 +	50.00	84.68	5.0e+03	4460	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-171-945-47 +	49.00	107.34	272.87	420
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-210-361-1 -	50.00	84.68	5.0e+03	4460	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-532-803-5 +	49.00	107.32	273.76	421
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-089-986-1 +	50.00	83.82	5.6e+03	4837	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-211-631-12 +	49.00	107.32	273.76	421
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-478-585-1 +	50.00	83.82	5.6e+03	4837	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-265-628-12 +	49.00	107.32	273.76	421
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-717-312-1 +	50.00	83.82	5.6e+03	4837	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-141-10 +	49.00	107.32	273.76	421
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-266-408-1 +	50.00	83.82	5.6e+03	4837	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-653-403-13 +	49.00	107.32	273.76	421
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-446-855A-1 +	50.00	83.82	5.6e+03	4837	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-44 -	49.00	102.99	476.84	632
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-991-408-1 +	50.00	83.16	6.1e+03	5145	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-171-945-18 +	49.00	100.96	618.98	765
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-403-545-1 +	50.00	81.31	7.7e+03	6122	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-171-945-56 +	49.00	100.96	618.98	765
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-404-381-1 +	50.00	81.31	7.7e+03	6122	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-860-882A-28 +	49.00	100.79	632.28	777
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-015-973-2 +	50.00	80.00	9.0e+03	6924	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-936-165A-118 -	49.00	98.92	803.51	926
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-448-164-2 +	50.00	80.00	9.0e+03	6924	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-365-150-4 -	49.00	98.37	862.14	975
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-150-741-1 +	50.00	77.30	1.3e+04	8920	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-975-526-4 +	49.00	97.74	935.42	1035
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-150-741-1 +	50.00	77.30	1.3e+04	8920	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-07-974-409C-426 +	49.00	97.74	935.42	1035
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-073-492-1 +	50.00	71.44	2.7e+04	15462	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-457-046B-57 +	49.00	95.17	1.4e+03	1317
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-125-468-1 -	50.00	64.38	6.5e+04	30001	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-853-559A-27 +	49.00	94.58	1.4e+03	1392
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-474-933-1 -	50.00	64.38	6.5e+04	30001	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-487-550-7 -	49.00	93.85	1.5e+03	1437
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-336-198C-2 -	49.50	99.86	712.77	954	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-949-637-1 -	49.00	93.85	1.5e+03	1491
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-886-765-1 +	49.50	99.54	742.53	983	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-416-488-1 -	49.00	92.22	1.9e+03	1737
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-115-660-1 +	49.50	99.54	742.53	983	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-421-050A-7 +	49.00	92.22	1.9e+03	1737
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-179-557-14 +	49.50	97.55	958.49	1185	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-664-800-7 +	49.00	92.22	1.9e+03	1737
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-924-440-1 +	49.50	96.62	1.1e+03	1293	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-665-309-7 +	49.00	92.22	1.9e+03	1737
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-56 -	49.50	95.76	1.2e+03	1402	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-661-569-7 +	49.00	92.22	1.9e+03	1737
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-54 -	49.50	95.76	1.2e+03	1402	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-737-226-3 +	49.00	91.33	2.1e+03	1858
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-121-425-4 +	49.50	95.57	1.2e+03	1426	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-826-246-1 +	49.00	90.97	2.2e+03	1953
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-583-562B-11 +	49.50	95.08	1.3e+03	1494	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-944-495-1 +	49.00	90.97	2.2e+03	1953
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-779-113-11 +	49.50	95.08	1.3e+03	1494	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-136-640-1 +	49.00	90.97	2.2e+03	1953
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-815-667-8 +	49.50	94.91	1.3e+03	1518	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-925-588-1 -	49.00	90.97	2.2e+03	1953
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-546-829A-1 -	49.50	94.53	1.4e+03	1573	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-288-292A-1 +	49.00	89.89	2.6e+03	2163
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-876-365-1 -	49.50	94.53	1.4e+03	1573	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-381-849-2 +	49.00	88.81	2.9e+03	2393
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-230-388-2 +	49.50	94.52	1.4e+03	1575	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-073-383-1 +	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-07-721-761A-27 +	49.50	93.88	1.5e+03	1672	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-073-383-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-07-978-687-27 +	49.50	93.88	1.5e+03	1672	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-328-239A-1 +	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-05801-05801-27 +	49.50	93.14	1.7e+03	1792	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-328-239A-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-965-762-7 +	49.50	92.83	1.8e+03	1845	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-328-239A-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-07-978-687-28 +	49.50	92.83	1.8e+03	1845	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-328-239A-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-07-978-687-28 +	49.50	92.83	1.8e+03	1845	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-328-239A-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-392-806A-1 +	49.50	92.34	1.9e+03	1932	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-328-239A-1 -	49.00	88.69	3.0e+03	2420
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-257-490-1 +	49.50	92.34	1.9e+03	1932	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-354-129-9 +	49.00	88.32	3.1e+03	2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-350-741-1 +	49.50	92.08	1.9e+03	1980	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-121-714-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-463-875A-1 +	49.50	92.08	1.9e+03	1980	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-322-742-13 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-965-903B-7 -	49.50	91.28	2.1e+03	2135	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-477-108A-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-553-619B-6 +	49.50	91.09	2.2e+03	2173	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-477-108A-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-714-918-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-477-112-1 +	49.00	87.99	3.3e+03	2584
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-265-315-84 -	49.50	89.49	2.7e+03	2525	1					



/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-485-568A-3 + 49.00 84.58 5.1e+03 3561 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-357-698-5 + 49.00 84.58 5.1e+03 3561 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-590-554A-3 + 49.00 84.58 5.1e+03 3561 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-184-223-3 + 49.00 84.58 5.1e+03 3561 1  
/cgn2\_6/ptodata/2/ina/PCBUS\_COMB.seq:PCT-US93-12682-5 + 49.00 84.58 5.1e+03 3561 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-040-774-1 - 49.00 82.60 6.5e+03 4285 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-233-788A-39 + 49.00 82.31 6.8e+03 4406 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-038-682-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-302-832-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-530-198-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-880-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-728-470-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-617-697-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-719-641-3 + 49.00 81.10 7.9e+03 4937 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-468-036-23 + 49.00 80.76 8.2e+03 5093 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-376-843-23 + 49.00 80.76 8.2e+03 5093 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-310-912A-157 - 49.00 80.68 8.3e+03 5134 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-301-085-157 - 49.00 80.68 8.3e+03 5134 1  
/cgn2\_6/ptodata/2/ina/PCBUS\_COMB.seq:PCT-US95-04589-157 - 49.00 80.68 8.3e+03 5134 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-329-920-1 + 49.00 80.31 8.7e+03 5313 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-875-154-2 + 49.00 78.72 1.1e+04 6169 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-324-867-1 + 49.00 77.32 1.3e+04 7032 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-480-604A-9 + 49.00 77.22 1.3e+04 7101 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-403-496A-9 + 49.00 77.22 1.3e+04 7101 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-915-136-9 + 49.00 77.22 1.3e+04 7101 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-542-003-1 + 49.00 74.87 1.7e+04 8855 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-322-760A-1 + 49.00 74.87 1.7e+04 8855 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-038-682-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-302-832-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-530-198-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-880-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-728-470-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-617-697-6 + 49.00 74.32 1.9e+04 9233 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-719-641-6 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-920-812-2 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-920-812-2 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-921-177-2 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-362-578-2 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-920-812-2 + 49.00 73.36 2.1e+04 10207 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-080-044-1 + 49.00 72.76 2.3e+04 10803 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-222-617A-26 + 49.00 72.14 2.5e+04 11444 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-222-617A-26 + 49.00 71.32 2.7e+04 12364 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-144-085-3 + 49.00 60.69 1.0e+05 3529 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-770-379-17 - 49.00 60.20 1.1e+05 35100 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-757-669A-17 - 49.00 60.20 1.1e+05 35100 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-230-371A-17 - 49.00 60.20 1.1e+05 35100 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-090-793-12 + 49.00 58.78 1.3e+05 40138 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-471-119A-1 + 49.00 57.12 1.5e+05 46899 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-920-422-17 - 49.00 56.66 1.6e+05 48974 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:5191064-4 + 48.50 109.82 198.65 296 1 Patent  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-681-812-1 + 48.50 101.47 579.32 648 1 S  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-353-133-1 - 48.50 99.29 765.97 795 1 S  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-955-629C-1 + 48.50 99.15 780.49 806 1 S  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-366-316-4 + 48.50 98.00 904.67 898 1 S  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-936-165A-148 + 48.50 97.75 933.69 919 1 S  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-475-316A-107 - 48.50 97.52 961.55 939 1 S  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-181-271A-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-449-315-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-444-803-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-449-043-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-456-265A-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-455-416-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-455-244-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-454-876-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-457-364-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-456-262-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-456-240-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-455-730-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-971-217-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-350-600-26 + 48.50 94.90 1.3e+03 1201 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-448-744-5 - 48.50 94.14 1.5e+03 1290 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-876-885-25 - 48.50 93.53 1.6e+03 1366 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-257-581-4 + 48.50 92.41 1.9e+03 1518 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-257-581-6 + 48.50 92.41 1.9e+03 1518 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-486-721A-2 + 48.50 92.14 1.9e+03 1556 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-001-826-24 + 48.50 91.52 2.1e+03 1650 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-142-623-12 + 48.50 91.23 2.2e+03 1695 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-353-133-2 - 48.50 90.50 2.4e+03 1815 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-663-566A-16 + 48.50 90.32 2.4e+03 1846 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-023-610-16 + 48.50 90.32 2.4e+03 1846 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-288-065A-16 + 48.50 90.32 2.4e+03 1846 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-362-240A-16 + 48.50 90.32 2.4e+03 1846 1  
/cgn2\_6/ptodata/2/ina/PCBUS\_COMB.seq:PCT-US95-10245-16 + 48.50 90.32 2.4e+03 1846 1  
/cgn2\_6/ptodata/2/ina/PCBUS\_COMB.seq:PCT-US94-09752-2 + 48.50 90.12 2.5e+03 1846 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-201-31 + 48.50 90.04 2.5e+03 1896 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-197-31 + 48.50 90.04 2.5e+03 1896 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-201-31 + 48.50 90.04 2.5e+03 1896 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-041-991A-9 + 48.50 90.00 2.5e+03 1902 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-041-991A-9 + 48.50 90.00 2.5e+03 1902 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-714-918-44 + 48.50 89.49 2.7e+03 1996 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-265-315-44 + 48.50 89.49 2.7e+03 1996 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-265-315-44 + 48.50 89.49 2.7e+03 1996 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-266-417-44 + 48.50 89.49 2.7e+03 1996 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-382-106-1 - 48.50 89.48 2.7e+03 1998 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-471-033-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-471-033-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-046A-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-470-566B-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-470-566B-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-334-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-334-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-300-529-6 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-300-529-18 + 48.50 89.45 2.7e+03 2004 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-07-779-890-1 + 48.50 88.69 3.0e+03 2152 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-07-779-890-1 + 48.50 88.69 3.0e+03 2152 1  
/cgn2\_6/ptodata/2/ina/PCBUS\_COMB.seq:PCT-US93-035640-1 + 48.50 88.34 3.1e+03 2223 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-257-073-4 + 48.50 88.34 3.1e+03 2223 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-191-8 + 48.50 87.98 3.3e+03 2301 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-200-8 + 48.50 87.98 3.3e+03 2301 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-232-197-8 + 48.50 87.98 3.3e+03 2301 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-836-567-1 + 48.50 87.97 3.3e+03 2303 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-468-558-4 + 48.50 87.61 3.4e+03 2382 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-676-444-4 + 48.50 87.61 3.4e+03 2382 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-861-464-11 + 48.50 87.28 3.6e+03 2457 1  
/cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-396-001-11 + 48.50 87.28 3.6e+03 2457 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-323-433A-11 + 48.50 87.28 3.6e+03 2457 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-471-033-35 + 48.50 86.77 3.8e+03 2576 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-35 + 48.50 86.77 3.8e+03 2576 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-35 + 48.50 86.77 3.8e+03 2576 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-470-566B-35 + 48.50 86.77 3.8e+03 2576 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-300-529-35 + 48.50 86.77 3.8e+03 2576 1  
/cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-471-033-4 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-471-033-17 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-4 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-17 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-044-26 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-4 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-17 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-463-483A-26 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-046A-4 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-046A-17 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-471-046A-26 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-470-566B-17 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-470-566B-26 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-334-4 + 48.50 86.45 4.0e+03 2655 1  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-469-334-17 + 48.50 86.45 4.0e+03 2655 1



```
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-26 + 48.50 86.45 4.0e+03 2655
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-300-529-4 + 48.50 86.45 4.0e+03 2655
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-300-529-17 + 48.50 86.45 4.0e+03 2655
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-300-529-26 + 48.50 86.45 4.0e+03 2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-773-608A-1 + 48.50 86.44 4.0e+03 2658
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-232-200-44 + 48.50 86.23 4.1e+03 2710
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-232-200-70 + 48.50 86.23 4.1e+03 2710
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-232-197-70 + 48.50 86.23 4.1e+03 2710
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-232-201-44 + 48.50 86.23 4.1e+03 2710
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-232-201-70 + 48.50 86.23 4.1e+03 2710

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-216-276A-30

seq_documentation_block:
; Sequence 30, Application US/08216276A
; Patent No. 5595912
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
; TITLE OF INVENTION: ASSOCIATED WITH US IBDV VARIANTS, VECTOR CARRYING DNA
; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID
; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,276A
; FILING DATE: 23-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/083,784
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/519,202
; FILING DATE: 04-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/227,311
; FILING DATE: 02-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1622 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: GLS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-08-216-276A-30
```

```
alignment_scores:
Quality: 77.00 Length: 125
Ratio: 1.305 Gaps: 6
Percent Similarity: 47.200 Percent Identity: 21.600

alignment_block:
US-09-528-682-3 x US-08-216-276A-30 ..
Align seg 1/1 to: US-08-216-276A-30 from: 1 to: 1622

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
::: ||||| ::|||
1 ATTCCAACCAACGAGATAACCCAG.....CCAAAT 29

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetG 49
: ::::: :::::
30 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAAGTGGTGC 73

49 InAspGlyAsp.....LeuPheGlyThr 56
||:::|||
74 AGAAGGGGACCATGTCATGTCGGCAAGTGGGAGCCTAGCAGTGACG 123

57 ValAsnGlnSerAsnPhePro..... 63
:::
124 ATTCATGTTGGCAACTATCCAGGGGCCCTCCGTCGCTCACACTAGTAGC 173

64 .....MetSerT 66

174 CTACGAAAGAGTGGCAACAGGATCTGCTTACGTCGCTGGGTGAGCA 223

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
::|||
224 ACTTCGAGCTGATCCCAATCTGTAAGTACCAAGAACCTGGTTACAGAA 273

83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
:::
274 TACGGCCGATTTGACCCAGGAGCCATGAAGTACACAAAATTGATA...CT 320

96 eMetGluAspAspPheIleGlyVal 104
: |||
321 GAGTGAGAGGAGCCGCTTGGCATC 345

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-216-276A-18

seq_documentation_block:
; Sequence 18, Application US/08216276A
; Patent No. 5595912
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID
; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
; TITLE OF INVENTION: ASSOCIATED WITH US IBDV VARIANTS, VECTOR CARRYING DNA
; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID
; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,276A
; FILING DATE: 23-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/083,784
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/519,202
; FILING DATE: 04-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/227,311
; FILING DATE: 02-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1622 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: GLS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-08-216-276A-30
```



```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/083,784
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/519,202
; FILING DATE: 04-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/227,311
; FILING DATE: 02-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3230 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: GLS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..3149
; US-08-216-276A-18

alignment_scores:
  Quality: 77.00 Length: 125
  Ratio: 1.305 Gaps: 6
  Percent Similarity: 47.200 Percent Identity: 21.600

alignment_block:
  US-09-528-682-3 x US-08-216-276A-18 ..

Align seg 1/1 to: US-08-216-276A-18 from: 1 to: 3230
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
999 ATCCAAACCAACGAGATACCCAG.....CCAAAT 1027
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
1028 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAGAGTGGTGC 1071
49 InAspGlyAsp.....LeuPheGlyThr 56
1072 AGGAAGGGGACCAAGATGTCATGTCGGCGAAGTGGAGCGCTCAGCAGTGC 1121
57 ValAsnGlnSerAsnPhePro..... 63
1122 ATTCATGTTGGCAACTATCCAGGGGCCCTCCGTCGTCACACTAGTAGC 1171
64 .....MetSerT 66
1172 CTACGAAAGAGTGGCAACAGGATCTGCTACGTCGCTGGGTGAGCA 1221
66 hPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
1222 ACTTCGAGCTGATCCCAATCTGAACTAGCAAGAAACCTGGTTACAGAA 1271
83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
1272 TACGGCCGATTTGACCCAGGAGCCATGAACTACACAAATGATA...CT 1318
96 eMetGluAspPheIleGlyVal 104

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-219-262B-11

1319 GAGTGAGAGGACCGCCTTGGCATC 1343

seq_documentation_block:
; Sequence 11, Application US/08219262B
; Patent No. 5788970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3230 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..3149
; US-08-219-262B-11

alignment_scores:
  Quality: 77.00 Length: 125
  Ratio: 1.305 Gaps: 6
  Percent Similarity: 47.200 Percent Identity: 21.600

alignment_block:
  US-09-528-682-3 x US-08-219-262B-11 ..

Align seg 1/1 to: US-08-219-262B-11 from: 1 to: 3230
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
999 ATCCAAACCAACGAGATACCCAG.....CCAAAT 1027
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
1028 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAGAGTGGTGC 1071
49 InAspGlyAsp.....LeuPheGlyThr 56
1072 AGGAAGGGGACCAAGATGTCATGTCGGCGAAGTGGAGCGCTCAGCAGTGC 1121
57 ValAsnGlnSerAsnPhePro..... 63
1122 ATTCATGTTGGCAACTATCCAGGGGCCCTCCGTCGTCACACTAGTAGC 1171
64 .....MetSerT 66
1172 CTACGAAAGAGTGGCAACAGGATCTGCTACGTCGCTGGGTGAGCA 1221
66 hPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
1222 ACTTCGAGCTGATCCCAATCTGAACTAGCAAGAAACCTGGTTACAGAA 1271
83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
1272 TACGGCCGATTTGACCCAGGAGCCATGAACTACACAAATGATA...CT 1318
96 eMetGluAspPheIleGlyVal 104
```















```

96 eMetGluAspAppheileglyVal 104
:   |||   |||   |||   |||   |||
1334 GAGTGAGAGGGACCGCTTGGCATC 1358

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-708-541A-27

seq_documentation_block:
; Sequence 27, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708.541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3261 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..531
US-08-708-541A-27

```

```

alignment_scores:
  Quality: 73.00      Length: 125
  Ratio: 1.259      Gaps: 6
Percent Similarity: 46.400      Percent Identity: 21.600

alignment_block:
US-09-528-682-3 x US-08-708-541A-27  ..
Align seg 1/1 to: US-08-708-541A-27 from: 1 to: 3261

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
:   |||   |||   |||   |||   |||
1016 ATTCCACAAACAGAGATACCCAG.....CCAAAT 1044

32 uAspGlyArgTy-SerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
:   |||   |||   |||   |||   |||
1045 CATATCC.....ATCAAACTGGAGATAGTACCTCCAAAGTGGTGC 1088

49 lnAspGlyAsp.....LeuPheGlyThr 56
:   |||   |||   |||   |||
1089 AGCAGGGGATCAGATGTCTATGTCGCGCAAGAGGGACCTAGCAGTGACG 1138

```

```

57 ValAsnGlnSerAsnPhePro..... 63
:   |||   |||   |||   |||   |||
1139 ATCCATGGTGGCAACTATCCAGGGGCCCTCCGTCCTCAGCTAGTAGTGGC 1188

64 .....MetSert 66
:   |||   |||   |||   |||   |||
1189 CTAGAAAGAGTGGCAACAGGATCCGTCGTACGGTCGCTGGGTGAGCA 1238

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
:   |||   |||   |||   |||   |||
1239 ACTCGAGCTGATCCCAATCTGAACTAGCAAGAACCTGGTTACAGAA 1288

83 AsnValIlyTyrrAsp.....MetAsnPhelLysLysLeuLeuArgph 96
:   |||   |||   |||   |||   |||
1289 TAGCGCGCATTTACCCAGGAGCATGAATACACAAAATTGATA...CT 1335

96 eMetGluAspAppheileglyVal 104
:   |||   |||   |||   |||   |||
1336 GAGTGAGAGGGACCGCTTGGCATC 1360

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-708-541A-29

seq_documentation_block:
; Sequence 29, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708.541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3261 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 131..3166
US-08-708-541A-29

```

```

alignment_scores:
  Quality: 73.00      Length: 125
  Ratio: 1.259      Gaps: 6
Percent Similarity: 46.400      Percent Identity: 21.600

```











```

TITLE OF INVENTION: HVT-050 and Uses Thereof
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,065A
FILING DATE: Aug-09-94
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0526
TELEX: 422523
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3350 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 129...2522
88-288-065A-1

nnment_scores:
Quality: 73.00 Length: 125
Ratio: 1.259 Gaps: 6
Percent Similarity: 46.400 Percent Identity: 21.600

nnment_block:
09-528-682-3 x US-08-288-065A-1 ..
ign seg l/1 to: US-08-288-065A-1 from: 1 to: 3350

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
::: ||||| :::|||||
1011 ATTCACAACAAAGAGATACCAG.....CCAAAT 1039

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
: : ::::: ::::::::::::::|
1040 CACATCC.....ATCAAACCTGGAGATAGTGACCTCCAAAAGTGTGGTC 1083

49 lnAspGlyVsp.....LeuPheGlyThr 56
|| ||||||
1084 AGGCAGGGGGATCAGATGTTATGTGCGCAAGAGGGGCCCTAGCAGTGACG 1133

57 ValAsnGlnSerAsnPhePro..... 63
::::: ::|||::|
1134 ATCCATGTTGGCACTATCCAGGGGCCCTCGGTCCCCTCACGTAGTAGGC 1183

64 .....MetSert 66
.....::|::|

1184 CTACGAAGAAGTGCCAACAGGATCCGTCGTACGTCGCTGGGTGAGCA 1233

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
::|::|::| ::|||::| ::|
1234 ACTTCGAGCTGATCCCAATCCTGAAGTAGCAAGAAGAACCTGGTTACAGAA 1283

```



```
1040 CACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAGTGGTGGTC 1083
49  InAspGlyAsp.....LeuPheGlyThr 56
||| |||||
1084 AGCGAGGGATCAGATGTTATGTCGCGAAGAGGGAGCCTAGCAGTGACG 1133
57  ValAsnGlnSerAsnPhePro..... 63
||||| :|||:||||
1134 ATCCATGTTGGCAACTATCCAGGGGCCCTCCGTCGCGTCACGCTAGTGCG 1183
64 .....MetSert 66
1184 CTACGAAAGAGTGGCAACAGGATCCGTCGTACGGTCTGCGGTGAGCA 1233
66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
:||||| :||||| :|||:|||||
1234 ACTTCGAGCTGATCCCAATCTCGAAGTACGAAAGAACCTGTTACAGAA 1283
83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
:||||| :||||| :||||| :|||:
1284 TACGGCCGATTGACCCAGGAGCCATGACACAAATTTGATA...CT 1330
96 eMetGluAspPheIleGlyVal 104
: ||| ||| :|||:
1331 GAGTGAGAGGACCGTCTTGGCATC 1355
```

seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: PCT-US95-10245-1

```
seq_documentation_block:
; Sequence 1, Application PC/TUS9510245
; GENERAL INFORMATION:
; APPLICANT: SYNTRO CORPORATION
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10245
; FILING DATE: 09-AUG-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 129..2522
PCT-US95-10245-1

alignment_scores:
```

```
Quality: 73.00 Length: 125
Ratio: 1.259 Gaps: 6
Percent Similarity: 46.400 Percent Identity: 21.600

alignment_block:
US-09-528-682-3 x PCT-US95-10245-1 ..
Align seg 1/1 to: PCT-US95-10245-1 from: 1 to: 3350

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
::| ||||| :|||:|||||
1011 ATTCAACAACAGAGATAACCCAG.....CCAAT 1039

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetG 49
:||||| :||||| :|||:|||||
1040 CACATC.....ATCAAACTGGAGATAGTGACCTCCAAAGTGGTGGTC 1083

49 InAspGlyAsp.....LeuPheGlyThr 56
||| |||||
1084 AGCGAGGGATCAGATGTTATGTCGCGAAGAGGGAGCCTAGCAGTGACG 1133

57 ValAsnGlnSerAsnPhePro..... 63
||||| :|||:|||||
1134 ATCCATGTTGGCAACTATCCAGGGGCCCTCCGTCGCGTCACGCTAGTGCG 1183

64 .....MetSert 66
1184 CTACGAAAGAGTGGCAACAGGATCCGTCGTACGGTCTGCGGTGAGCA 1233

66 hrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 82
:||||| :||||| :|||:|||||
1234 ACTTCGAGCTGATCCCAATCTCGAAGTACGAAAGAACCTGTTACAGAA 1283

83 AsnValLysTyrAsp.....MetAsnPheLysLysLeuLeuArgPh 96
:||||| :||||| :||||| :|||:
1284 TACGGCCGATTGACCCAGGAGCCATGACACAAATTTGATA...CT 1330

96 eMetGluAspPheIleGlyVal 104
: ||| ||| :|||:
1331 GAGTGAGAGGACCGTCTTGGCATC 1355
```

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: US-08-530-492-2

```
seq_documentation_block:
; Sequence 2, Application US/08530492
; Patent No. 5689052
; GENERAL INFORMATION:
; APPLICANT: Brown, Sherri M.
; APPLICANT: Dean, Duff A.
; APPLICANT: Fromm, Michael E.
; APPLICANT: Sanders, Patricia R.
; TITLE OF INVENTION: Synthetic DNA Sequences Having Enhanced
; TITLE OF INVENTION: Expression in Monocotyledonous Plants and Method For
; TITLE OF INVENTION: Preparation Thereof
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B&F
; STREET: 700 Chesterfield Parkway No. 5689052th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,492
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/172,333
```







;; TYPE: DNA  
;; ORGANISM: Saccharomyces cerevisiae  
US-09-356-952-11

## alignment\_scores:

Quality: 70.00 Length: 101  
Ratio: 1.250 Gaps: 4  
Percent Similarity: 55.446 Percent Identity: 25.743

## alignment\_block:

US-09-528-682-3 x US-09-356-952-11 ..

Align seg 1/1 to: US-09-356-952-11 from: 1 to: 5398

21 ValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArg..... 35

2044 GTAGATAGTGTACATGATCATGTTTCAGTTCCATTGTCGGTAAACGTAA 2093

36 .....TyrSerAsnPheA 40

2094 TGTAACTACCACTACACCGATACATTGACTCCAAATGAGATCATCATTC 2143

40 laLeuIleSerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThr 56

2144 GTACAGTCAAT...GAGACGATATGGAATAATTCTCAGTCTTAGCTCCA 2190

57 ValAsnGlnSerAsn.....PheProMetSe 65

2191 AGAAATAGTGTGTAATTCTGCTGTAACCAAGGACTTCAATACAAAATTC 2240

65 rThrPheGluGlnVal...ProAsnAsnLysGluPheLysGlyValIle 81

2241 TACTTTGGAAGATTTTTCACCGTCCACCAAAAATTTTAAAGTCAGCTAAAT 2290

81 exAlaAsnValLysTyrAspMetAsnPhelLysLysLeuLeuArgPheMet 97

2291 CGATTTACGAAATGGTGTGATGTTGGAATTCGGAATTTTAAAGCATGTT 2340

98 Glu 98

2341 CAG 2343

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-103-840A-2

## seq\_documentation\_block:

;; Sequence 2, Application US/09103840A  
;; Patent No. 6294328

## GENERAL INFORMATION:

;; APPLICANT: FLEISCHMAN, Robert D.

;; APPLICANT: WHITE, Owen R.

;; APPLICANT: FRASER, Claire M.

;; APPLICANT: VENTER, John C.

;; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

;; FILE REFERENCE: 24366-20007.00

;; CURRENT APPLICATION NUMBER: US/09/103,840A

;; CURRENT FILING DATE: 1998-06-24

;; NUMBER OF SEQ ID NOS: 2

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 2

;; LENGTH: 4403765

;; ORGANISM: Mycobacterium tuberculosis

;; FEATURE:

;; OTHER INFORMATION: CDC 1551

;; OTHER INFORMATION: "n" bases at various positions throughout the sequence

;; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

## alignment\_scores:

Quality: 70.00 Length: 75

Ratio: 1.489 Gaps: 4  
Percent Similarity: 62.667 Percent Identity: 28.000

## alignment\_block:

US-09-528-682-3 x US-09-103-840A-2/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-2 from: 1 to: 4403765

9 GlnAlaTyrGluProIleGluValAsnThrAsnThrValThrGlnIleAs 25

4044856CGCGGTACAAAGCCGTTGATACCTCGGTG...GCGATAACACAGATCAA 4044812

25 nGlySerAsn.....GluValProLeuAspGlyArgTyrSerAsnPheA 40

4044811CGCGGACACGTCACAGCGCACAGATCGGCGACGACAGAGTGC 4044762

40 laLeuIleSerAlaGluGlyMetGln...AspGly..... 51

4044761GGGCTGATCCTGAAGAAGGGTAAACAACGACGACCGGTCGAGAGGTC 4044712

52 .....AspLeuPheGlyThrValAs 58

4044711ATCACCAGTACCCACCGGCTACGCGCTGACCTGTTCACGCGCTCAG 4044662

58 nGlnSerAsnPheProMetSerThr 66

4044661CGGCCAAAAACGCGAAGTCCAGCAG 4044637

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-103-840A-1

## seq\_documentation\_block:

;; Sequence 1, Application US/09103840A  
;; Patent No. 6294328

## GENERAL INFORMATION:

;; APPLICANT: FLEISCHMAN, Robert D.

;; APPLICANT: WHITE, Owen R.

;; APPLICANT: FRASER, Claire M.

;; APPLICANT: VENTER, John C.

;; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

;; FILE REFERENCE: 24366-20007.00

;; CURRENT APPLICATION NUMBER: US/09/103,840A

;; CURRENT FILING DATE: 1998-06-24

;; NUMBER OF SEQ ID NOS: 2

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 1

;; LENGTH: 4411529

;; TYPE: DNA

;; ORGANISM: Mycobacterium tuberculosis

;; OTHER INFORMATION: H37Rv

US-09-103-840A-1

## alignment\_scores:

Quality: 70.00 Length: 75  
Ratio: 1.489 Gaps: 4  
Percent Similarity: 62.667 Percent Identity: 28.000

## alignment\_block:

US-09-528-682-3 x US-09-103-840A-1/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-1 from: 1 to: 4411529

9 GlnAlaTyrGluProIleGluValAsnThrAsnThrValThrGlnIleAs 25

4052790CGCGGTACAAAGCCGTTGATACCTCGGTG...GCGATAACACAGATCAA 4052744

25 nGlySerAsn.....GluValProLeuAspGlyArgTyrSerAsnPheA 40

4052743CGCGGACACGTCACAGCGCACAGATCGGCGACGACAGAGTGC 4052694

40 laLeuIleSerAlaGluGlyMetGln...AspGly..... 51



4052693GGCTGATCCTGTAAGAAGGTTAAACAAGAGACGCGGTCGAGAGGTC 4052644

52 .....AspLeuPheGlyThrValAs 58  
|||||

4052643ATCACCAAGTACCCACCGGTACCGCTGACCTGTTCAACGGCTCAG 4052594

58 nGlnSerAsnPheProMetSerThr 66  
: |||||

4052593CGCCAAAACGCGAAGTCAAGCAG 4052569

seq\_name: /cqn2\_6/ptodata/2/1na/5A\_COMB.seq:US-07-944-525-1

seq\_documentation\_block:

```
; Sequence 1, Application US/07944525
; Patent No. 5632989
; GENERAL INFORMATION:
; APPLICANT: SNYDER, DAVID B.
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: LUETTICKEN, HEINRICH D.
; TITLE OF INVENTION: ATTENUATED, LIVE VACCINE FOR DELAWARE
; TITLE OF INVENTION: STRAIN IBDV
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER AND
; STREET: 1755 JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22202
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/944,525
; FILING DATE: 19920914
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KELBER, STEVEN B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2284-028-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3180 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..3099
; US-07-944-525-1
```

```
alignment_scores:
  Quality: 69.00      Length: 123
  Ratio: 1.190      Gaps: 6
Percent Similarity: 47.154      Percent Identity: 21.951
```

alignment\_block:

```
US-09-528-682-3 x US-07-944-525-1
Align seg 1/1 to: US-07-944-525-1 from: 1 to: 3180
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
::: ||||||| :|||
949 ATTCCAACCAATGATGATAACCCAG.....CCAAT 977
```

```
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly.....G 47
: : : : : : : : : : : : : : : : : : : : : : : : : : :
978 CACATCCATCATACTGGAGATAGTACCTCCAAAAGTGGTGCAGGCAG 1027
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 lyMetGln.....AspGlyAspLeuPheGlyThrValAsn 58
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1028 GGGACAGATGTCTATGGTGGCAAGTGGGAGCCTAGCAGTGCAGATCCAT 1077
59 GlnSerAsnPhePro..... 63
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1078 GGTGGCAACTATCCAGGAGCCCTCCGTCACACTAGTGGCCTACGA 1127
64 .....MetSerThrPheG 68
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1128 AAGAGTGGCAACAGGATCTGCTTACGGTCCGCTGGGTGAGCAACTCG 1177
68 luGlnValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnVal 84
|| : : : : : : : : : : : : : : : : : : : : : : : : : :
1178 AGCTGATCCCAATCCTGAACTAGCAAGAACCTGGTTACAGAATACGGC 1227
85 LysTyrAsp.....MetAsnPheLysLysLeuLeuArgPheMetG1 98
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1228 CGATTGACCCAGGAGCCATGAACACGAAATTGATA...CTGAGTGA 1274
98 uAspAspPheIleGlyVal 104
| ||| : : : : : : : : : : : : : : : : : : : : :
1275 GAGGAGACCACCTTGGCATC 1293
```

seq\_name: /cqn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-219-262B-13

seq\_documentation\_block:

```
; Sequence 13, Application US/08219262B
; Patent No. 5788970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; APPLICANT: MENGEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
```



```
;
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..3099
US-08-219-262B-13

alignment_scores:
  Quality: 69.00      Length: 123
  Ratio: 1.190       Gaps: 6
  Percent Similarity: 47.154      Percent Identity: 21.951

alignment_block:
US-09-528-682-3 x US-08-219-262B-13
..
Align seg 1/1 to: US-08-219-262B-13 from: 1 to: 3180

16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
   ::: ||||| ::|||
949 ATTCCAAACCAATGAGATAACCCAG.....CCAAAT 977

32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly.....G 47
   ::: ::|||
978 CACATCCATCATCTACTGGAGATAGTACCTCCAAAAGTGATGGTCAGGCAG 1027

47 lyMetGln.....AspGlyAspLeuPheGlyThrValAsn 58
   |||
1028 GGGAAACAGATGTCATGTCGGCAAGTGGGAGCCTAGCAGTGCAGTCCAT 1077

59 GlnSerAsnPhePro..... 63
   ::|||
1078 GGTGGCAACTATCCAGGAGCCCTCCCGTCGCTCACACTAGTGGCCTACGA 1127

64 .....MetSerThrPheG 68
1128 AGAGTGGCAACGAGATCTCTGTTACGCTCGCTGGGGTGAGCAACTTCG 1177

68 luGlnValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnVal 84
   || ::|||
1178 AGCTGATCCCAATCTGTAAGTACGAGCAAGAACCTGTTACAGATACGGC 1227

85 LysTyrAsp.....MetAsnPheLysLysLeuLeuArgPheMetG1 98
   ::|||
1228 CGATTTCAGCCAGGAGCATGTAAGTACAGCAAAATTGATA...CTGAGTGA 1274

98 uAspAspPheIleGlyVal 104
   |||
1275 GAGGACCACTTGGGCATC 1293

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-031-655-13

seq_documentation_block:
; Sequence 13, Application US/09031655
; Patent No. 6017759
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; APPLICANT: MENGEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```



```

; Sequence 31, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..531
; US-08-708-541A-31

alignment_scores:
  Quality: 69.00 Length: 117
  Ratio: 1.232 Gaps: 5
  Percent Similarity: 47.863 Percent Identity: 24.786

alignment_block:
US-09-528-682-3 x US-08-708-541A-31 ..
Align seg 1/1 to: US-08-708-541A-31 from: 1 to: 3264

16 ValAsnThrAsnThrValThrGln..... 23
1019 GTCCCAACAATGATGATCCACCACCCATCATCTCCATGAACCTAGAGGT 1068
24 .....TleAsnGlySerAsnGluValProLeuAspGlyArgT 36
1069 TGTGACCTACAAAGATTGGCGCGACCGTGTGACCAATATCATGGACAG 1118
36 yrSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspGlyAsp 52
1119 TGAGTGGTACACTAGCTGTGACGTTGCACGGAGGCAACTACCCGCGGCT 1168
53 Leu.....PheGlyThrVa 57
1169 CTCGCTCTGTCAACCTGGTGGCTATGACACGAGTGGCTGCAGGATCTGT 1218
57 lAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAsnAsnL 74

```

```

1219 T...GTCACAGTTGCAGGGGTGAGCAACTTCGAGCTAATCCCAACCCCTG 1265
74 ysGluPheLysGlyValIleSerAlaAsnValLysTyAsp..... 87
1266 AGCTTGCAGAAACCTAGTTACAGAGTATGCCGCTTTGACCCCGGAGCA 1315
88 MetAsnPheLysLysLeuLeuArgPheMetGluAspAspPheIleGlyVa 104
1316 ATGAACCTACACCAACTAATA...CTGAGTGAGAGAGATCGTCTAGGCAT 1362
104 1 104
1363 C 1363

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-708-541A-33

seq_documentation_block:
; Sequence 33, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 131..3169
; US-08-708-541A-33

alignment_scores:
  Quality: 69.00 Length: 117
  Ratio: 1.232 Gaps: 5
  Percent Similarity: 47.863 Percent Identity: 24.786

alignment_block:
US-09-528-682-3 x US-08-708-541A-33 ..
Align seg 1/1 to: US-08-708-541A-33 from: 1 to: 3264

16 ValAsnThrAsnThrValThrGln..... 23

```



```

1019 GTCCCAACAATGAGATCACCACGCCCATCTTCCATGAACTAGAGGT 1068
      ||| ||||| :::|||||
24 .....IleAsnGlySerAsnGluValProLeuAspGlyArgT 36
      |||:::|||||::: |||:::|
1069 TGTGACCTACAAGATTGGCGCACCAGCTGGTGACCCCAATATCATGGACAG 1118
      |||:::|||||::: |||:::|
36 yfSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAsp 52
      |||::: |||:::|
1119 TGAAGTGGTACACTAGCTGTGACGGTGCACGGAGCAACTACCCCTGGGGCT 1168
      |||::: |||:::|
53 Leu.....PheGlyThrVa 57
      |||
1169 CTCCGTCCTGTCACTCCCTGGTGCCCTATGAACGAGCTGCAGGATCGT 1218
      |||
57 IasGlnSerAsnPheProMetSerThrPheGluGlnValProAsnAsnL 74
      |||::: |||:::|
1219 T...CTCAGCTTGCAGGGGTGAGCAACTTCGAGCTAATCCCCAACCCCTG 1265
      |||::: |||:::|
74 ysGluPheLysGlyValIleSerAlaAsnValLysTrpAsp..... 87
      :: |||:::|||||::: |||:::|
1266 AGCTTGCAAAGNACCTAGTTACAGAGTATGCGCCCTTGGACCCCGGAGCA 1315
      :: |||:::|||||::: |||:::|
88 MetAsnPheLysLysLeuArgPheMetGluAspPheIleGlyVa 104
      |||::: |||:::|
1316 ATGAACATACACCAACTAATA...CTGAGTGCAGAGAGATCGTCTAGGCAT 1362
      |||::: |||:::|
104 1 104
      :
1363 C 1363

```

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:US-08-910-501-3

```

seq_documentation_block:
; Sequence 3, Application US/08910501
; Patent No. 6020159
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin
; APPLICANT: Lonetto, Michael
; APPLICANT: Warren, Patrick
; TITLE OF INVENTION: NOVEL 3'-DEHYDROQUINATE SY
; TITLE OF INVENTION: NTHASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.501
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: P50549-1
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 base pairs

```

[illegible]

seq\_name: /can2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-398-550-3

```

seq_documentation_block:
; Sequence 3, Application US/09398550
; Patent No. 6232292
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin
; Lonetto, Michael
; Warren, Patrick
; TITLE OF INVENTION: NOVEL 3-DEHYDROQUINATE SY
; NTHASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,550
; FILING DATE: 17-Sep-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:

```











```
seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:US-09-379-523-4
seq_documentation_block:
; Sequence 4, Application US/09379523
; Patent No. 6280993
```



Wed Jun 19 08:58:17 2002

```

2936 GACATTATCTGTTTATATAAACAGCATTCGTCTCAAAATATATCCCTAT 2985
      87 AspMetAsnPhelysLysLeu 93
          ::::|||||:
2986 TCTTTAAACATAAAGGATTA 3006

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq.US-08-998-416-30

seq_documentation_block:
; Sequence 30, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Stehner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Melgs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1007RP
US-08-998-416-30

alignment_scores:
Quality: 65.50 Length: 63
Ratio: 1.489 Gaps: 4
Percent Similarity: 69.841 Percent Identity: 31.746

alignment_block:
US-09-528-682-3 x US-08-998-416-30/rev ..
Align seg 1/1 to reverse of: US-08-998-416-30 from: 1 to: 518

53 LeuPheGlyThrValAsnGlnSerAsnPhePrometSerThrPheClu1 69
||||| |||||:|||||:
222 CTATTCACCACCAACCATCGGTTCTCTCCACAGACTAAGGACATTGG 173

GENERAL INFORMATION:
; APPLICANT: YAMATO, Ichiro
; APPLICANT: HOSAKA, Yoshiaki
; TITLE OF INVENTION: GENE ENCODING CLASS I COLLAGENASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch disk
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: WINDOWS 98
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/379,523
; FILING DATE: August 24, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wayne, Milton J.
; REGISTRATION NUMBER: 17,906
; REFERENCE/DOCKET NUMBER: U-Wp-5462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-683-8150
; TELEFAX: 212-532-4285
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3356
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-379-523-4

alignment_scores:
Quality: 66.50 Length: 107
Ratio: 1.090 Gaps: 3
Percent Similarity: 57.009 Percent Identity: 20.561

alignment_block:
US-09-528-682-3 x US-09-379-523-4 ..
Align seg 1/1 to: US-09-379-523-4 from: 1 to: 3356

7 LeuGlnGlnAlaTyrluPuIlleGluValAsnThrAsnThrValThrGl 23
:|||||:|||||:
2689 ATAAAGAGCGTATGTGTCCTCAATAGTTGAAGTGTTACTGTAAAGGTGA 2738
:|||||:|||||:
23 nileAsnGlySerAsn..... 28
:|||||:|||||:
2739 TTAAATGGTCTGCATGATGCTGATACCTCTCTATTTTGATGTAAGAAG 2788
29 .....GluValProLeuAspGlyArgTyrSerAsnPhe 39
|||||:|||||:
2789 ATGTGTGATGTTACAATTCAAACCTCTTATTTCAGGGTCA...TCTAATTTC 2835
|||||:|||||:
40 AlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyTh 56
:|||||:|||||:
2836 ACATGGTTAGTTTAAAGAGGAGAGCATCAAAACCATATATGCAAGTGG 2885
56 rValAsnGlnSerAsnPhePrometSerThrPheGluGlnValProAsnA 73
:|||||:|||||:
2886 TATAGATAAGAATAACTCAAAGTTGGAACATTTAAATCTACAAAAGGAA 2935
73 snLyGlu.....PheLysGlyValIleSerAlaAsnValLysTyr 86
:|||||:|||||:

```



```
69 nValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLysT 86
   :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
172 GATCAACAGAACAGCATGCAAGGGCGCATTTCCACGGAG...GAGT 126
   :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

86 yf.....AspMetAsnPheLysLysLeuLeuArgPheMetGlu 98
   || :::::|||||:::|||||:::|||||:::|||||:::|||||
125 ACACCAACTTCATCAATCTCAGATTGAGGAAGTCATCCGCTTCAGGAG 76
   :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

99 AspAspPheIleGly.....ValHisGlyGlu 107
   :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
75 GAG.....GTCGGCTAGATGTCCTTGTTCATGGTGAA 43
```

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-936-165A-232

```
seq_documentation_block:
; Sequence 232, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 232:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-936-165A-232
```

alignment\_scores:  
Quality: 65.50 Length: 135  
Ratio: 1.008 Gaps: 5  
Percent Similarity: 48.148 Percent Identity: 19.259

```
alignment_block:
US-09-528-682-3 x US-08-936-165A-232
Align seg 1/1 to: US-08-936-165A-232 from: 1 to: 904
15 GluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVal... 30
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
160 AAGTCAATCGCTCACACATTAAATGAAGTTGAGAGTAACAATGAAGTTAA 209
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
30 ..... 30
210 TTTTCAGCAAGAGAAATTAAANGNACCGGCTAGTATATTGAATAAATCAG 259
31 .....ProLeuAspGlyArgTyrSerAsnPheAla 40
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
260 GTAAAGGTATAAAATGAGACAAATTACAGGGCGGATCAAAAGACTTTTGT 309
41 .....LeuIleSerAlaGluGly...Me 48
   |||:::|||||:::|||||:::|||||:::|||||:::|||||
310 TCTAAAGCTAAGCGGATTCGGATTTCACCCACATCAGATGGTGAATAAT 359
48 tGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetS 65
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
360 ATTGATCGGAACATTAGTCAATCGGTAAAGGTCAGTGACAATAACTTAA 409
65 erThrPheGluGlnValProAsnAsnLysGluPheLysGly..... 78
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
410 AAGATTGAAAGATATTGTTAAACAAACGTGTGTTCAAGGTTTCAGATAAT 459
79 .....ValIleSerAlaAsnValLysTyrAspMet..... 88
   ::|||:::|||||:::|||||:::|||||:::|||||:::|||||
460 CAATTAAGTGAATTAATTCGTGTTCTGTAATAATACGACAAACGTTATGA 509
89 .....AsnPheLysLysLeuLeuArgPheMetGluA 99
   |||:::|||||:::|||||:::|||||:::|||||:::|||||
510 TAATTTCCCGATTTTGAACAATAGTAAGGCGATGTTGAACTTTAATATAG 559
99 spasp 100
   ::|||
560 AAGAT 564
```

seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:PCT-US93-05944-1

```
seq_documentation_block:
; Sequence 1, Application PC/TUS9305944
; GENERAL INFORMATION:
; APPLICANT: Lin et al., Hun-Chi
; TITLE OF INVENTION: Molecular cloning of the genes
; TITLE OF INVENTION: responsible for collagenase product
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harris Brotman
; STREET: 401 B. St Ste 1700
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05944
; FILING DATE: 19930622
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brotman, Harris F.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-3630
; TELEFAX: (619) 236-1048
; INFORMATION FOR SEQ ID NO: 1:
```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2817 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2808
; PCT-US93-05944-1

alignment_scores:
    Quality: 65.50      Length: 107
    Ratio: 1.092        Gaps: 3
    Percent Similarity: 56.075    Percent Identity: 20.561

alignment_block:
    US-09-528-682-3 x PCT-US93-05944-1 ..
    Align seg 1/1 to: PCT-US93-05944-1 from: 1 to: 2817

7 LeuGlnGlnAlaTyGluProIleGluValAsnThrAsnThrValThrGl 23
2143 ATAAAGAGCGTCAATGGTCCCAATAGTTGAAGGTGTACTGTAAGAGGTGA 2192
23 nIleAsnGlySerAsn..... 28
2193 TTTAANTGGTCTGATGATGCTGNTACCTTCATTATTTGATGTAAGAGAG 2242
29 .....GluValProLeuAspGlyArgTyrSerAsnPhe 39
2243 ATGGTGATGTTACAATTGAACCTTCCTTATTCAGGGTCA...TCTAATTC 2289
40 AlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyTh 56
2290 ACATGGTTAGTTTATAAAGAGGGACGACATCAAAATCATATTGCAAGTGG 2339
56 rValAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAsnA 73
2340 TATAGATAGAATAACTCAAAAGTTGGAAACATTTAAAGCTACAAAAGGAA 2389
73 snLysGlu.....PheLysGlyValIleSerAlaAsnValLysTyr 86
2390 GACATTATGCTGTTATATATAAACACGAGGATTCGCTTCAATATATCTTAT 2439
87 AspMetAsnPheLysLysLeu 93
2440 TCTTTAAACATATAAAGGATTA 2460

```

seq\_name: /cqn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-676-967-4

```

seq_documentation_block:
; Sequence 4, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM pc compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,967

```

```

;
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36, 627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-676-967-4

alignment_scores:
    Quality: 65.00      Length: 116
    Ratio: 1.140        Gaps: 4
    Percent Similarity: 49.138    Percent Identity: 24.138

alignment_block:
    US-09-528-682-3 x US-08-676-967-4/rev ..
    Align seg 1/1 to reverse of: US-08-676-967-4 from: 1 to: 2277

18 ThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspG1 34
   |||:|||:||||:||||:||||:||||:||||:||||:||||:||||:
1353 ACCAGCGGGATCACACCTTCACGAGCCAGGTACAGTTACGGGTACCGG 1304
   34 yArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspG 51
   |||:|||:||||:||||:||||:||||:||||:||||:||||:||||:
1303 TCGGTTTTTAACTTT.....GGTGTCTGCAGTTTA 1272

51 lyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPhe 67
   |||:|||:||||:||||:||||:||||:||||:||||:||||:||||:
1271 GCAGTCTCGTCAGCGGTACAGCCAGGTCACTTTCAGTGCAGCACCGTC 1222

68 GluGlnValProAsnAsnLysGluPheLys..... 77
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1221 CAGTTTCAGACCACCAAGCTTCGTTTTCGGGGAAGCAGCAGCAGGCATT 1172

78 .....GlyV 79
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1171 TCTGAGCAGCTTCTCGGTCATGAACCTGAGCGAAGCGCAACCTTTGGAG 1122

79 alIleSerAlaAsnVallystYrAspMetAsn.....PheLysLys 92
   |||:|||:||||:||||:||||:||||:||||:||||:||||:||||:
1121 TGTTCGGGTCCGGGTGCAGACGATACGACGATATTCAGTTTCAACCGAA 1072

93 LeuLeu...ArgPheMetGluAspPheIleGlyValHisGlyGlu 107
   |||:|||:||||:||||:||||:||||:||||:||||:||||:||||:
1071 CTGCTGCAGCAGCTTCACCCAGTTCTTCTTCGGAGTCGAAGGACAG 1024

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-676-974-4

seq_documentation_block:
; Sequence 4, Application information US/08676974
; Patent No. 5770422
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104

```



```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,974
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-676-974-4

alignment_scores:
  Quality: 65.00 Length: 116
  Ratio: 1.140 Gaps: 4
  Percent Similarity: 49.138 Percent Identity: 24.138

alignment_block:
US-09-528-682-3 x US-08-676-974-4/rev ..
Align seg 1/1 to reverse of: US-08-676-974-4 from: 1 to: 2277

18 ThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspG1 34
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1353 ACCAGCAGGATCAGACCTTCACGAGCCAGGTACAGGTACGGGTACCGG 1304
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
34 yArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspG 51
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1303 TCGGTTTTTTAACTTT.....GGTGTCTGCAGTTTA 1272

51 lYAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPhe 67
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1271 GCAGTTCCTCAGGGTAAACAGCCAGGTCACTTTTCAGCTGACGACGTC 1222
::: : : : : : : : : : : : : : : : : : : : : : : : : :
68 GluGlnValProAsnAsnLysGluPheLys.....GlyV 79
::: : : : : : : : : : : : : : : : : : : : : : : : : :
1221 CAGTTTCAGACCAGCAGCTTCGTTTTCGGGGAAGCAGCAGGCATT 1172
78 .....GlyV 79

1171 TCTGAGCAGCTTCCTGGGTGTCATGAACCTGAGCGAAAGCGCAACCTTTGGAG 1122
79 alileSerAlaAsnValLysTyrAspMetAsn.....PheLysLys 92
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1121 TGTTCGGTGTCCGGGTGTCAGACGATACGACGATATTTTCAGTTACCCGAA 1072

93 LeuLeu...ArgPheMetGluAspAspPheIleGlyValHisGlyGlu 107
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1071 CTGCTGCAGCAGTTTACCCAGTCTCTTCTTCTTCGGGAGTGAAGGACAG 1024

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-098-487-4

seq_documentation_block:
; Sequence 4, Application US/09098487
; Patent No. 5917025
; GENERAL INFORMATION:
; APPLICANT: COLLINS, Kathleen
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 11

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,487
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-098-487-4

alignment_scores:
  Quality: 65.00 Length: 116
  Ratio: 1.140 Gaps: 4
  Percent Similarity: 49.138 Percent Identity: 24.138

alignment_block:
US-09-528-682-3 x US-09-098-487-4/rev ..
Align seg 1/1 to reverse of: US-09-098-487-4 from: 1 to: 2277

18 ThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspG1 34
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1353 ACCAGCAGGATCAGACCTTCACGAGCCAGGTACAGGTACGGGTACCGG 1304
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
34 yArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspG 51
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1303 TCGGTTTTTTAACTTT.....GGTGTCTGCAGTTTA 1272

51 lYAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPhe 67
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1271 GCAGTTCCTCAGGGTAAACAGCCAGGTCACTTTTCAGCTGACGACGTC 1222
::: : : : : : : : : : : : : : : : : : : : : : : : : :
68 GluGlnValProAsnAsnLysGluPheLys.....GlyV 79
::: : : : : : : : : : : : : : : : : : : : : : : : : :
1221 CAGTTTCAGACCAGCAGCTTCGTTTTCGGGGAAGCAGCAGGCATT 1172
78 .....GlyV 79

1171 TCTGAGCAGCTTCCTGGGTGTCATGAACCTGAGCGAAAGCGCAACCTTTGGAG 1122
79 alileSerAlaAsnValLysTyrAspMetAsn.....PheLysLys 92
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1121 TGTTCGGTGTCCGGGTGTCAGACGATACGACGATATTTTCAGTTACCCGAA 1072

93 LeuLeu...ArgPheMetGluAspAspPheIleGlyValHisGlyGlu 107
||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1071 CTGCTGCAGCAGTTTACCCAGTCTCTTCTTCTTCGGGAGTGAAGGACAG 1024

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-058-489-34

```



```
seq_documentation_block:
; Sequence 34, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-086A
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 10091
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-34

alignment_scores:
  Quality: 63.00      Length: 73
  Ratio: 1.465        Gaps: 2
Percent Similarity: 58.904 Percent Identity: 28.767

alignment_block:
US-09-528-682-3 x US-09-058-489-34/rev ..

Align seg 1/1 to reverse of: US-09-058-489-34 from: 1 to: 10091

      2 PhePheThrArgAlaLeuGlnGlnAlaTyr.....ValThrG 23
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      8000 TTCTGCACCTAGGCACCTCCAGAGAGGATTTCAGAGACGCGATTGTGATACAT 7951

      12 ....GluProIleGluValAsnThrAsnThr.....ValThrG 23
      : : : : : : : : : : : : : : : : : : : : : : : :
      7950 TAAAGAGGACATTATGATGTAACCAACCAACCTACATTTTGTGTGACGG 7901

      23 IntLeuAsnGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPhe 39
      : : ||||| : : ||||| : : ||||| : : ||||| : : |||||
      7900 AGAAGAACGCACAGTCATCATACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 7851

      40 AlaLeuIleSerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyTh 56
      ||| : : : : : : : : : : : : : : : : : : : : : :
      7850 TTCTTTGGTGTGAATCCAGTGTGTAAGAGGAATCTAGCAGCAAGCTGAA 7801

      56 rValAsnGlnSerAsnPhe 62
      : : : : : : : : : : : : : : : : : : : : : :
      7800 TACTATCATAGTAATTC 7782

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-802-141-3

seq_documentation_block:
; Sequence 3, Application US/08802141
; Patent No. 5773009
; GENERAL INFORMATION:
; APPLICANT: GLASS, ROGER I.
; APPLICANT: GENTSCH, JOHN R.
; APPLICANT: BHAN, M. K.
; APPLICANT: DAS, BIMAL K.
; TITLE OF INVENTION: ROTAVIRUS STRAIN AND RELATED
; TITLE OF INVENTION: COMPOSITIONS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,141
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/231,041
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2352 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..2329
US-08-802-141-3

alignment_scores:
  Quality: 62.50      Length: 105
  Ratio: 1.008        Gaps: 6
Percent Similarity: 59.048 Percent Identity: 25.714

alignment_block:
US-09-528-682-3 x US-08-802-141-3 ..

Align seg 1/1 to: US-08-802-141-3 from: 1 to: 2352

      8 GlnGlnAlaTyrGluProIleGluValAsnThrAsnThrValThrGlnIl 24
      ||||| ||| ||||| : : : : : : : : : : : : : :
      134 CAAAGTCATACGCACCGATTTCGTGGGGATCAGGAGAG...ACGTTAAG 180

      24 eAsnGlySerAsnGluValProLeuAspGlyArgTyr..... 36
      : : : : : : : : : : ||||| ||||| |||||
      181 CGAAATGTTGAAGAGCAACCTTTAGACGGGCCATATACACAGATAAGT 230

      37 .....SerAsnPhe...AlaLeuIleSerAlaGluGlyGlyMet 48
      ||||| : : : : : : : : : : : : : : : : : :
      231 CAAATTTGCGCTCTACTATTGGTATTATTAATCAATCCATCA..... 271

      49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPhePrometSe 65
      : : ||||| : : ||||| : : ||||| : : ||||| : : |||||
      272 AATGATGCGTGGTGTCTCGGTAACGGATAACAGTACGCTTTGGATGTT 321

      65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIle.... 80
      ||||| : : ||||| : : : : : : : : : : : :
      322 TACTTATTAGTCTTACCAATACAGCTAGACTAGTGTGCTAGTAATG 371

      81 .....SerAlaAsnValLysTyrAspMetAsnPheLysLysLeu 93
      : : : : : : : : : : : : : : : : |||||
      372 TAATGATGAGACAGTGAATATATATCAATAGAC...AACTAGGTTCCGCA 418

      94 LeuArgPheMetGlu 98
      : : ||||| : : : : :
      419 TATAAATTTGTGGAT 433

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-718-575-3
seq_documentation_block:
```



```

; Sequence 3, Application US/07718575
; Patent No. 5202257
; GENERAL INFORMATION:
; APPLICANT: Heinemann Ph.D., Stephen F.
; APPLICANT: Boulter Ph.D., James R.
; APPLICANT: Hollmann Ph.D., Michael NMN
; APPLICANT: Bettler Ph.D., Bernhard NMN
; APPLICANT: Jensen Ph.D., Jan E.
; TITLE OF INVENTION: GLUTAMATE RECEPTOR COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: United States
; ZIP: 90071-2921
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/718,575
; FILING DATE: 19910813
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8962
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; TELEX: 9103330318
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3505 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: GluR2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 316..2967
; US-07-718-575-3

alignment_scores:
Quality: 62.50 Length: 64
Ratio: 1.894 Gaps: 3
Percent Similarity: 51.562 Percent Identity: 32.812

alignment_block:
US-09-528-682-3 x US-07-718-575-3 ..
Align seg 1/1 to: US-07-718-575-3 from: 1 to: 3505

42 lIeSerAlaGluGlyMetGlnAspGlyAspLeu.....Ph 54
||| ||||| ||||| ||||| ||||| |||||
1021 ATCATTCGCAATCTGGGATTCACATGCTGCTGAAATTCAGTT 1070

54 eGlyThrValAsnGlnSerAsnPhProMet.....64
||| ||||| ||||| ||||| |||||
1071 TGGAGGAGCAAAATCTCTGATTCAGATTGTAGACTAGCAGCATTC 1120

65 .....SerThrPheGluGlnValProAsn 72
||| ||||| ||||| ||||| |||||
1121 TGGTGTCTAAATTTATGAAGATGGTCAACACTGGAAGAG..... 1161

73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyr 86

```

```

; Sequence 3, Application US/07718575
; Patent No. 5202257
; GENERAL INFORMATION:
; APPLICANT: Heinemann Ph.D., Stephen F.
; APPLICANT: Boulter Ph.D., James R.
; APPLICANT: Hollmann Ph.D., Michael NMN
; APPLICANT: Bettler Ph.D., Bernhard NMN
; APPLICANT: Jensen Ph.D., Jan E.
; TITLE OF INVENTION: GLUTAMATE RECEPTOR COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: United States
; ZIP: 90071-2921
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/718,575
; FILING DATE: 19910813
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8962
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; TELEX: 9103330318
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3505 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: GluR2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 316..2967
; US-07-718-575-3

alignment_scores:
Quality: 62.50 Length: 64
Ratio: 1.894 Gaps: 3
Percent Similarity: 51.562 Percent Identity: 32.812

alignment_block:
US-09-528-682-3 x US-07-718-575-3 ..
Align seg 1/1 to: US-07-718-575-3 from: 1 to: 3505

42 lIeSerAlaGluGlyMetGlnAspGlyAspLeu.....Ph 54
||| ||||| ||||| ||||| ||||| |||||
1021 ATCATTCGCAATCTGGGATTCACATGCTGCTGAAATTCAGTT 1070

54 eGlyThrValAsnGlnSerAsnPhProMet.....64
||| ||||| ||||| ||||| |||||
1071 TGGAGGAGCAAAATCTCTGATTCAGATTGTAGACTAGCAGCATTC 1120

65 .....SerThrPheGluGlnValProAsn 72
||| ||||| ||||| ||||| |||||
1121 TGGTGTCTAAATTTATGAAGATGGTCAACACTGGAAGAG..... 1161

73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyr 86

```



```

54 eGlyThrValAsnGlnSerAsnPheProMet..... 64
1071 TGGAGGACAAATGCTCTGGATTAGATTAGACTACGACGATGCC 1120
65 .....SerThrPheGluGlnValProAsn 72
1121 TGGTGTCTAAATTTATAGAAAGATGGTCAACACTGGGAAGAG..... 1161

```

```

73 AsnLysGluPheIysGlyValIleSerAlaAsnValLysTyr 86
      |||||: |||: :|||: :|||: :|||:
1162 ..AAAGAAATACCTGGAGCACAACACAGCGACAATTAAGTAT 1200

```

seq\_name: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:US-08-486-269A-3

```

seq_documentation_block:
: Sequence 3, Application documentation US/08486269A
: Patent No. 5945509
: GENERAL INFORMATION:
: APPLICANT: Heinemann, Stephen F.
: APPLICANT: Boulter, James R.
: APPLICANT: Hollmann, Michael
: APPLICANT: Bettler, Bernhard
: APPLICANT: Jensen, Jan E.
: TITLE OF INVENTION: GLUTAMATE RECEPTOR COMPOSITIONS
: TITLE OF INVENTION: AND METHODS
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Gray Cary Ware & Freidenrich LLP
: STREET: 4365 Executive Drive, Suite 1600
: CITY: San Diego
: STATE: CA
: COUNTRY: USA
: ZIP: 92121
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows DEMONSTRATION Version 2.0.D
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/486,269A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/013,767
: FILING DATE: 04-FEB-1993
: APPLICATION NUMBER: 07/718,575
: FILING DATE: 21-JUN-1991
: APPLICATION NUMBER: PCT/US90/06153
: FILING DATE: 25-OCT-1990
: APPLICATION NUMBER: 07/428,116
: FILING DATE: 27-OCT-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Reiter, Stephen E
: REGISTRATION NUMBER: 31,192
: REFERENCE/DOCKET NUMBER: P41 9986
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-677-1409
: TELEFAX: 619-677-1465
: TELEX:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3505 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: Glur2
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 316...2964
: OTHER INFORMATION:
: US-08-486-269A-3

```

```

alignment_scores:
  Quality: 62.50      Length: 64
  Ratio: 1.894       Gaps: 3
  Percent Similarity: 51.562  Percent Identity: 32.812

alignment_block:
  US-09-528-682-3 x US-08-486-269A-3  ..

  Align seg 1/1  to: US-08-486-269A-3  from: 1  to: 3505

      42  ILeSerAlaGluGlyMetGlnAspGlyAspLeu.....Ph 54
      |||  |||:::  |||:::  ||| ||| ||| |||
      1021 ATCATGTGCAAACTGGGATTCTACTGATGGGACCTGCTGAAAATTTCAGTT 1070
      54  eGlyThrValasnGlnSerAsnPhProMet..... 64
      |||  |||:::  |||:::  ||| ||| ||| |||
      1071 TGGAGGAGCAAAATGCTCTGGATTTCAGATTGTAGACTACGACGATTCCTC 1120
      65  .....SerThrPheGluGlnValProAsn 72
      ||| ||| ||| ||| ||| ||| ||| |||
      1121 TGGTGCTCTAAATTATAGAAGATGCTCAACACTGGAAGAG..... 1161

      73  AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyr 86
      ||| |||:::  |||:::  ||| ||| ||| ||| |||
      1162 ...AAAGAATACCTCGGACGACACACAGCGCAATTAAGTAT 1200

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-397-602A-8

seq_documentation_block:
; Sequence 8, Application US/083977602A
; Patent No. 5646044
; GENERAL INFORMATION:
; APPLICANT: Wilson, Charles R
; APPLICANT: Tang, Maria R
; APPLICANT: Berger, Harald
; APPLICANT: Christianson, Teresa M
; APPLICANT: Hansen, Dieter
; TITLE OF INVENTION: Expression Systems for the Production
; TITLE OF INVENTION: of Target Proteins in Bacillus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Henkel Corporation Law Department
; STREET: 140 Germantown Pike, Suite 150
; CITY: Plymouth Meeting
; STATE: PA
; COUNTRY: USA
; ZIP: 19462
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,602A
; FILING DATE: 02-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jaeschke, Wayne C
; REGISTRATION NUMBER: 21,062
; REFERENCE/DOCKET NUMBER: D8969/M4828
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 832-2200
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA sequence of the
; DESCRIPTION: alkaline protease gene of strain ATCC 53928"

```



```
; DESCRIPTION: elements"
; ORIGINAL SOURCE:
; ORGANISM: Bacillus licheniformis
; STRAIN: ATCC 53926
; US-08-397-602A-8

alignment_scores:
  Quality: 62.00      Length: 42
  Ratio: 2.296        Gaps: 0
  Percent Similarity: 64.286  Percent Identity: 35.714

alignment_block:
  US-09-528-682-3 x US-08-397-602A-8  ..
  Align seg 1/1 to: US-08-397-602A-8  from: 1 to: 1452

  14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  1188 ATTGAACGGAACGTCATATGCTTCTCTCATGTAGCGGAGCAGCGTT 1237

  30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  1238 TGATCTTGTCAAACATCCGAACCTTTCAGCTTCACAAGTCCGCAACCGT 1287

  47 lyMetGlnAspGlyAspLeuPheGly 55
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  1288 CTCCTCCAGCAGCGGCACTATTATGGG 1313

seq_name: /cgn2_6/ptodata/2/lna/3A_COMB.seq:US-08-685-774-1

seq_documentation_block:
; Sequence 1, Application US/08685774
; Patent No. 5712147
; GENERAL INFORMATION:
; APPLICANT: Shih, Jason C. H.
; APPLICANT: Lin, Xiang
; APPLICANT: Miller, Eric S.
; TITLE OF INVENTION: DNA ENCODING BACILLUS LICHENIFORMIS
; TITLE OF INVENTION: PWD-1 KERATINASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5712147th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,774
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/250,028
; FILING DATE: 27-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5051-260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 420-2200
; TELEFAX: (919) 881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1457 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus licheniformis
; STRAIN: PWD-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 215..1354
; US-08-685-774-1

alignment_scores:
  Quality: 62.00      Length: 42
  Ratio: 2.296        Gaps: 0
  Percent Similarity: 64.286  Percent Identity: 35.714

alignment_block:
  US-09-528-682-3 x US-08-685-774-1  ..
  Align seg 1/1 to: US-08-685-774-1  from: 1 to: 1457

  14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  1174 ATTGAACGGAACGTCATATGCTTCTCTCATGTAGCGGAGCAGCGTT 1223

  30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  1224 TGATCTTGTCAAACATCCGAACCTTTCAGCTTCACAAGTCCGCAACCGT 1273

  47 lyMetGlnAspGlyAspLeuPheGly 55
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  1274 CTCCTCCAGCAGCGGCACTATTATGGG 1299

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-282-147-36

seq_documentation_block:
; Sequence 36, Application US/09282147
; Patent No. 6274147
; GENERAL INFORMATION:
; APPLICANT: YAKHARIA, Vikram
; APPLICANT: YAO, Kun
; TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
; TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
; TITLE OF INVENTION: TRANSCRIPTS
; FILE REFERENCE: 8288-9023
; CURRENT APPLICATION NUMBER: US/09/282,147
; CURRENT FILING DATE: 1999-03-31
; EARLIER APPLICATION NUMBER: US/60/080,278
; EARLIER FILING DATE: 1998-03-31
; EARLIER APPLICATION NUMBER: PCT/US97/12955
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Infectious pancreatic necrosis virus
; US-09-282-147-36

alignment_scores:
  Quality: 61.50      Length: 116
  Ratio: 1.008        Gaps: 4
  Percent Similarity: 52.586  Percent Identity: 18.966

alignment_block:
  US-09-528-682-3 x US-09-282-147-36  ..
  Align seg 1/1 to: US-09-282-147-36  from: 1 to: 3097

  3 PheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAs 19
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```







```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/521,872
; FILING DATE: 31-AUG-1995
; CLASSIFICATION: 436
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4849 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 148..4407
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; US-08-521-872-13

alignment_scores:
  Quality: 61.50      Length: 100
  Ratio: 1.500       Gaps: 3
  Percent Similarity: 41.000   Percent Identity: 26.000

alignment_block:
US-09-528-682-3 x US-08-521-872-13  ..

Align seg 1/1 to: US-08-521-872-13  from: 1 to: 4849

14  ileGluValAsnThrAsnThrValThrGluIleAsnGlySer..... 27
    :::::||||| :::::||||| :::::||||| :::::|||||
1438 CTGGAGGCAAAATGGTAGACTCGCAAGGTAATAATTCGTGAGCAA 1487
    28 .....AsnGluValProLeuAspGlyA 35
    1488 GACAGGACGCTAGACACACTTCTATTAATAAGAGGGAACACTCGGAACA 1537
    35 rqtYrSerAsnPheAlaLeuIleSerAlaGlu..... 45
    1538 GAGACAGAACCAAAATCTGCCAAGTGATTAAGAAGTCACTATATGGTAGAC 1587
    46 .....GlyClyMetGlnAspGlyAAspLeuPheClyThrValAs 58
    1588 AAGGAATGTGTTGGTCAGGATGAGGATGATTTATTTGGCGATAGCAA 1637
    58 nGlnSerAsn..... 61
    1638 TAATTCGAATTCACAAACCAATCGNACAAAGATATATCGGACCAAAATTA 1687
    62 .....PheProMetSerThrPheGluGlnValProAsnAsnLys 74
    1688 CCGAGGATATGTTCCAAATGCTCTGTGATGAAGAGAAAATAATAACAATAA 1737

```



















[illegible]



[illegible]



[illegible]



/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-20943	64.00	144.74	130.26	3
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15089	64.00	144.27	138.36	3
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-873-402A-7283	64.00	140.95	211.98	48
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-209-830-7283	64.00	140.95	211.98	48
/cgn2_6/ptodata/2/pna/US0619_COMB.seq:US-60-197-872-39564	64.00	140.79	216.44	52
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-021-323-14770	64.00	140.12	235.59	5
/cgn2_6/ptodata/2/pna/US0625_COMB.seq:US-60-255-619-14770	64.00	140.12	235.59	5
/cgn2_6/ptodata/2/pna/US0608_COMB.seq:US-60-081-131-21	64.00	139.83	244.71	538
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-669-817A-24436	64.00	139.58	252.75	6
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-873-402A-35787	64.00	138.87	276.61	5
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-209-830-35649	64.00	138.87	276.61	5
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-873-402A-3639	64.00	138.14	303.88	6
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-209-830-3639	64.00	138.14	303.88	6
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-708-427-61585	64.00	135.43	430.04	8
/cgn2_6/ptodata/2/pna/US0635_COMB.seq:US-60-356-057-2186	64.00	133.90	523.23	984
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-060-063-3980	64.00	133.48	552.17	102
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-614-150-23363	64.00	130.17	844.96	1
/cgn2_6/ptodata/2/pna/US0617_COMB.seq:US-60-173-464-19261	64.00	130.17	844.96	1
/cgn2_6/ptodata/2/pna/US0619_COMB.seq:US-60-191-637-23446	64.00	130.17	844.96	1
/cgn2_6/ptodata/2/pna/US0619_COMB.seq:US-60-191-681-18502	64.00	130.17	844.96	1
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-391-631-1266	64.00	128.36	1.1e+03	173
/cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-679-686A-17	64.00	128.24	1.1e+03	175
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-248-796-3871	64.00	127.31	1.2e+03	192
/cgn2_6/ptodata/2/pna/US0609_COMB.seq:US-60-096-409-3871	64.00	127.31	1.2e+03	192
/cgn2_6/ptodata/2/pna/US0631_COMB.seq:US-60-312-544-4460	64.00	126.27	1.4e+03	21
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-439A-20086	64.00	126.16	1.4e+03	2
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-708-427-40391	64.00	123.07	2.1e+03	2966
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-760-482-33	64.00	121.84	2.5e+03	3
/cgn2_6/ptodata/2/pna/US0636_COMB.seq:US-60-360-207-34178	64.00	119.40	3.4e+03	431
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-017-116897	64.00	118.39	3.8e+03	21
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-116897	64.00	118.39	3.8e+03	21
/cgn2_6/ptodata/2/pna/US0621_COMB.seq:US-60-213-359-1741	64.00	117.23	4.4e+03	53
/cgn2_6/ptodata/2/pna/US0632_COMB.seq:US-60-324-185-24095	64.00	117.11	4.5e+03	3
/cgn2_6/ptodata/2/pna/US0606_COMB.seq:US-60-068-138-324	64.00	111.20	9.6e+03	994
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-200-391-59	64.00	101.07	3.5e+04	2790
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-200-391-60	64.00	101.07	3.5e+04	2790
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-760-457-504	64.00	99.81	4.1e+04	317
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-760-483-691	64.00	99.81	4.1e+04	317
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-760-485-1122	64.00	99.81	4.1e+04	317
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-764-905-27771	64.00	99.81	4.1e+04	317
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-092-399-27771	64.00	99.81	4.1e+04	317
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-202-420-18	64.00	99.50	4.3e+04	3276
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-16794	64.00	97.41	5.6e+04	4
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-264-74305	64.00	97.41	5.6e+04	4
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-7050	64.00	97.38	5.7e+04	40
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-534-859-935	64.00	91.08	1.3e+05	772
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-736-935	64.00	91.08	1.3e+05	772
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-534-859-226	64.00	90.12	1.4e+05	852
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-736-938	64.00	90.12	1.4e+05	852
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-534-859-398	64.00	89.02	1.7e+05	953
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-736-398	64.00	89.02	1.7e+05	953
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-736-1305	64.00	87.96	1.9e+05	10
/cgn2_6/ptodata/2/pna/US0994_COMB.seq:US-09-404-520-24536	63.50	144.39	136.37	29
/cgn2_6/ptodata/2/pna/US0992_COMB.seq:US-09-248-796-13992	63.50	140.57	222.40	44
/cgn2_6/ptodata/2/pna/US0609_COMB.seq:US-60-096-409-13992	63.50	140.57	222.40	44
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-134-001C-45	63.50	140.38	228.12	450
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-468-147-73	63.50	140.35	228.76	451
/cgn2_6/ptodata/2/pna/US0619_COMB.seq:US-60-197-872-13012	63.50	140.12	235.81	4
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-397-022-3860	63.50	138.44	292.33	548
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-808-383-3860	63.50	138.44	292.33	54
/cgn2_6/ptodata/2/pna/US0617_COMB.seq:US-60-172-519-4482	63.50	137.55	327.66	60
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-874-708A-81661	63.50	137.42	333.17	6
/cgn2_6/ptodata/2/pna/US0621_COMB.seq:US-60-211-750-80528	63.50	136.92	355.39	640
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-021-323-8930	63.50	136.92	355.39	64
/cgn2_6/ptodata/2/pna/US0625_COMB.seq:US-60-255-619-8930	63.50	136.92	355.39	64
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-021-323-8887	63.50	136.83	359.59	64
/cgn2_6/ptodata/2/pna/US0625_COMB.seq:US-60-255-619-8887	63.50	136.78	361.69	64
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-021-323-8800	63.50	136.78	361.69	64
/cgn2_6/ptodata/2/pna/US0625_COMB.seq:US-60-255-619-8800	63.50	136.78	361.69	64
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-439A-17916	63.50	136.77	362.39	64
/cgn2_6/ptodata/2/pna/US0620_COMB.seq:US-60-207-458-116063	63.50	136.77	362.39	64
/cgn2_6/ptodata/2/pna/US09632_COMB.seq:US-60-325-448-3938	63.50	136.17	391.40	3
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-680-598A-737	63.50	135.96	402.71	3
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-388446	63.50	135.89	405.12	3
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-388446	63.50	135.89	405.12	3
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-861-478-325	63.50	135.86	407.15	3
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-620-394B-3355	63.50	134.52	483.26	3
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-974-300A-1877	63.50	134.39	491.46	3
/cgn2_6/ptodata/2/pna/US0894_COMB.seq:US-08-487-032A-51	63.50	134.25	500.44	8
/cgn2_6/ptodata/2/pna/US088A_COMB.seq:US-08-561-469A-51	63.50	134.25	500.44	8
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-993-002A-9613	63.50	134.25	500.44	8
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-993-002A-9613	63.50	134.25	500.44	8
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-708-427-1336	63.50	130.27	833.67	1
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-648-119-81	63.50	129.73	894.07	1
/cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-517-537-682	63.50	129.06	973.34	1
/cgn2_6/ptodata/2/pna/US0936_COMB.seq:US-09-360-207-29629	63.50	128.41	1.1e+03	3
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-723-667-35	63.50	127.72	1.2e+03	3
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-822-835A-581	63.50	127.72	1.2e+03	3
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-252-991A-5686	63.50	127.54	1.2e+03	3
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-887-272A-2729	63.50	125.30	1.6e+03	30
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-379-653-35	63.50	121.65	2.5e+03	3
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-723-667-35	63.50	121.65	2.5e+03	3
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-583-110-1375	63.50	121.01	2.7e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	120.99	2.7e+03	3
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-708-427-1336	63.50	120.61	2.9e+03	3
/cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-950-083-177	63.50	120.61	2.9e+03	3
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-497-191-53	63.50	118.49	3.8e+03	41
/cgn2_6/ptodata/2/pna/US0906_COMB.seq:US-60-061-989-517	63.50	118.21	3.9e+03	41
/cgn2_6/ptodata/2/pna/US0621_COMB.seq:US-60-213-359-2145	63.50	117.74	4.2e+03	3
/cgn2_6/ptodata/2/pna/US0632_COMB.seq:US-60-324-185-28084	63.50	117.74	4.2e+03	3
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-252-991A-5730	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50	117.46	4.3e+03	3
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-107-433-2240	63.50			







```

/csn2_5/ptodata/2/pna/US6036_COMB.seq:US-60-360-207-25837 + 62.50 110.49 1.1e+04
/csn2_6/ptodata/2/pna/US094_COMB.seq:US-09-404-520-4970 + 62.50 107.90 1.5e+04
/csn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-20916 + 62.50 106.63 1.7e+04
/csn2_6/ptodata/2/pna/US087_COMB.seq:US-08-781-986A-171 + 62.50 106.48 1.8e+04
/csn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171-171 + 62.50 106.48 1.8e+04
/csn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171B-171 + 62.50 106.48 1.8e+04
/csn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171C-171 + 62.50 106.48 1.8e+04
/csn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171E-171 + 62.50 106.48 1.8e+04
/csn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-24789 + 62.50 102.08 3.1e+04
/csn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-038-081-371 + 62.50 101.19 3.5e+04
/csn2_6/ptodata/2/pna/US6004_COMB.seq:US-60-046-983-381 + 62.50 101.19 3.5e+04

```

```

seq_documentation_block:
; Sequence 1769, Application US/09522305
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Donovan, Michael J.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL LIBRARY
; FILE REFERENCE: 1600.1085-001
; CURRENT APPLICATION NUMBER: US/09/522,305
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/123,524
; EARLIER FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 1861
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1769
; LENGTH: 770
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-522-305-1769

```

```

alignment_scores:
  Quality: 80.50      Length: 129
  Ratio: 1.238       Gaps: 5
  Percent Similarity: 50.388      Percent Identity: 27.132

alignment_block:
US-09-528-682-3 x US-09-522-305-1769

```

Align seg 1/1 to: US-09-522-305-1769 from: 1 to: 770

1 AspPhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAs 17  
|||||  
197 GATTTCCTTCATTTCTGGACATCGCTTTGGGCCCGCATGATGCCAT 246  
17 nThrAsnThrValThr.....GlnIleAsnGlySerAsnG 29  
::: :: ||| ::::| |  
247 GGCCTGGGCTTATATCTATTTTGAATAATCAAGATCTTCCACTCCAACG 296  
29 luVal.....ProLeuAspGlyArgTyrSerAsnPhe 39  
|||| | ||| ::::| |  
297 AGGTTATGAGTCAGTTCTATCCACTCTTGATCGGAAACGGAGTGAGAA 346  
40 AlaLeuIle..... 42  
::: :::  
347 GTGTATGTGTCACCCAGAGGACAGTAGCCAGCTGCTGGAAGGAGATGAC 396  
43 .....SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyT 56  
||| ::| |::| ::| ::| ::|  
397 AAATGCTGCAGTAGTCAGAACGAGGCGGTACTAACCATGGAGCAGTTCACAA 446  
56 hrValAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAsn 72  
|||| ::| ::| ::| ::| ::| ::| ::| ::|  
447 CTGTCTCCTCAAGAGTACCTTCCTCTCAAGACAGAGAGCAAAATCCAGGAG 496  
73 AsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAsp...Me 88  
||| ||| ::::| ::| ::| ::|  
497 CTGATGAGGCGGCGGTGCATCTCCAGCAGCAGCAAGCAATGCGAGACTTGGT 546

```

88  tasnPheLysLysLeuLeuArgPheMetGluAspAsp 100
   :|||||:|||||  |||||||
547  CAACTACCGCTCACTG.....TTTATGAGGATGAC 577

seq_name: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-634-306B-203708

seq_documentation_block:
; Sequence 203708, Application US/09634306B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/09/634, 306B
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203708
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Human
US-09-634-306B-203708

```

```
alignment_scores:
  Quality: 78.00      Length: 113
  Ratio: 1.300       Gaps: 5
  Percent similarity: 53.097    Percent Identity: 25.664

alignment_block:
  US-09-538-682:3    x  US-09-634-306B-203708/rev  ..
```

```

Align seg 1/1 to reverse of: US-09-634-306B-203708 from: 1 to: 1194

2  PhePheThrArgAlaLeuGlnAlaTyrGluProfilGluValAsnth 18
      ||||| :|||:||||||| ||| :|||:|||||:
863 TTCTTCAGAGCCAGCATCCAACAAGCCACAGAAGCTCTGAGTATTTCCCT 814

18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
      :|||:|||||:|||||:||||| :||| :|||
813 TTCTTCAGTCACCCACATAAATGGCTTCAGGGCCCTTCCTGGGGAAGCCCT. 765

32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMet 48
      :|||:|||||:|||||:||||| :|||:|||||
764 ..GAAGGAGAGATTACAGCATACACTTGTGGYACATTGAAGC..... 723

49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
      ||| :||| :|||
722 .....TTCACTCTCTCC 712

65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerA 82
      ||| ||| :|||:||||| :|||:|||||
711 TCAAGGATCCAATCTCCCTCAGTCAAGAAGCTCCAGGTATCTGAACATG 662

82 la.....AsnValLysTyrAspMetAsnPheLysLysLeu... 93
      :||| :||| :|||
661 GATGCCAGGTCATAAATCCCCACTATGTGTACTCCATCAGGTCTCTGTCC 612

94 .....LeuArgPheMetGluAspAspPheIleGluVal 104

```



```
|||||.....
611 TCAGAACTAGAGCTAGTAAAGATAGACTCATGGGAGTC 573
seq_name: /cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-234-611-7401

seq_documentation_block:
; Sequence 7401, Application US/09234611
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-757
; CURRENT APPLICATION NUMBER: US/09/234,611
; CURRENT FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 21025
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7401
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-234-611-7401

alignment_scores:
  Quality: 77.50      Length: 80
  Ratio: 1.890        Gaps: 2
  Percent Similarity: 51.250  Percent Identity: 27.500

alignment_block:
US-09-528-682-3 x US-09-234-611-7401/rev ..
Align seg 1/1 to reverse of: US-09-234-611-7401 from: 1 to: 455

2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
||||| .....
454 TTCTTCAGAGCCAGCATCCACAGCCACAGAGCTGTGAGTATTCCT 405

18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProLeuA 33
:|||||.....
404 TTCCTCAGTCAGCCACACATAAATGCTTCAGGGCCTCTGGGGAAGGCTTG 355

33 spGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGln 49
:|||||.....
354 AAGGAGATTACAGCATACACTTGTGGCAGCATTTGAAGC..... 314

50 AspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerTh 66
||| ...
313 .....TTCACCTCTTCCTCA 300

66 rPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||| ||| .....
299 AGGATCCAATCTCCCTCAGTCAAGAAGCTCCAGGTATC 260

seq_name: /cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-235-076-11038

seq_documentation_block:
; Sequence 11038, Application US/09235076
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/235,076
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11038
; LENGTH: 455
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-235-076-11038

alignment_scores:
  Quality: 77.50      Length: 80
  Ratio: 1.890        Gaps: 2
  Percent Similarity: 51.250  Percent Identity: 27.500

alignment_block:
US-09-528-682-3 x US-09-235-076-11038/rev ..
Align seg 1/1 to reverse of: US-09-235-076-11038 from: 1 to: 455

2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
||||| .....
454 TTCTTCAGAGCCAGCATCCACAGCCACAGAGCTGTGAGTATTCCT 405

18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProLeuA 33
:|||||.....
404 TTCCTCAGTCAGCCACACATAAATGCTTCAGGGCCTCTGGGGAAGGCTTG 355

33 spGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGln 49
:|||||.....
354 AAGGAGATTACAGCATACACTTGTGGCAGCATTTGAAGC..... 314

50 AspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerTh 66
||| ...
313 .....TTCACCTCTTCCTCA 300

66 rPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||| ||| .....
299 AGGATCCAATCTCCCTCAGTCAAGAAGCTCCAGGTATC 260

seq_name: /cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-248-797-32436

seq_documentation_block:
; Sequence 32436, Application US/09248797
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-764
; CURRENT APPLICATION NUMBER: US/09/248,797
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 48909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32436
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-248-797-32436

alignment_scores:
  Quality: 77.50      Length: 80
  Ratio: 1.890        Gaps: 2
  Percent Similarity: 51.250  Percent Identity: 27.500

alignment_block:
US-09-528-682-3 x US-09-248-797-32436/rev ..
Align seg 1/1 to reverse of: US-09-248-797-32436 from: 1 to: 455

2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
```







```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-904-809-7401
```

```
alignment_scores:
  Quality: 77.50      Length: 80
  Ratio: 1.890        Gaps: 2
  Percent Similarity: 51.250  Percent Identity: 27.500
```

```
alignment_block:
```

```
US-09-528-682-3 x US-09-904-809-7401/rev ..
```

```
Align seg 1/1 to reverse of: US-09-904-809-7401 from: 1 to: 455
```

```
2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
454 TTCCTCAGCAGCCATCCACAGCCAGCAGAGCTGTGAGTATTCCT 405
```

```
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProLeuA 33
: : : : : ||||| : : : : : ||| : : : : :
404 TTCCTCAGTACCCACATAAATGGCTTCAGGCGCTTCTGGGGAAGGCTTG 355
```

```
33 spGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGln 49
: : : : : : : : : : : : : : : ||| : : : : :
354 AAGGAAGATTTACAGCATACACTTGTGGCAGCATTTGAAGGC..... 314
```

```
50 AspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerTh 66
||||| : : : : :
313 .....TTCACCTCTTCCTCA 300
```

```
66 rPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||||| : : : : : |||||
299 AGGGATCCAATCTCCCTCAGTCAGTCAAGAGCTCCAGGTATC 260
```

```
seq_name: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-925-564-32436
```

```
seq_documentation_block:
; Sequence 32436, Application US/09925564
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-764
; CURRENT APPLICATION NUMBER: US/09/925,564
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/248,797
; NUMBER OF SEQ ID NOS: 48909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32436
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-564-32436
```

```
alignment_scores:
  Quality: 77.50      Length: 80
  Ratio: 1.890        Gaps: 2
  Percent Similarity: 51.250  Percent Identity: 27.500
```

```
alignment_block:
```

```
US-09-528-682-3 x US-09-925-564-32436/rev ..
```

```
Align seg 1/1 to reverse of: US-09-925-564-32436 from: 1 to: 455
```

```
2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
454 TTCCTCAGCAGCCATCCACAGCCAGCAGAGCTGTGAGTATTCCT 405
```

```
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProLeuA 33
: : : : : ||||| : : : : : ||| : : : : :
404 TTCCTCAGTACCCACATAAATGGCTTCAGGCGCTTCTGGGGAAGGCTTG 355
```

```
33 spGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGln 49
: : : : : : : : : : : : : : : ||| : : : : :
354 AAGGAAGATTTACAGCATACACTTGTGGCAGCATTTGAAGGC..... 314
```

```
50 AspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerTh 66
||||| : : : : :
313 .....TTCACCTCTTCCTCA 300
```

```
66 rPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||||| : : : : : |||||
299 AGGGATCCAATCTCCCTCAGTCAGTCAAGAGCTCCAGGTATC 260
```

```
seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-477-574A-1
```

```
seq_documentation_block:
; Sequence 1, Application US/09477574A
; GENERAL INFORMATION:
; APPLICANT: Stram, Yehuda
; APPLICANT: Rogel, Arle
; APPLICANT: Sela, Ilan
; APPLICANT: Edelman, Orit
; APPLICANT: Shachar, Yehoshua
; APPLICANT: Zarnberg, Yehuda
; APPLICANT: Gontmakher, Tanya
; APPLICANT: Khayat, Eli
; TITLE OF INVENTION: RECOMBINANT VACCINES AGAINST IBDV
; FILE REFERENCE: 0231.00013
; CURRENT APPLICATION NUMBER: US/09/477,574A
; CURRENT FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/114,634
; PRIOR FILING DATE: 1999-01-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3054
; TYPE: DNA
; ORGANISM: Infectious bursal disease virus
US-09-477-574A-1
```

```
alignment_scores:
  Quality: 77.00      Length: 127
  Ratio: 1.283        Gaps: 6
  Percent Similarity: 47.244  Percent Identity: 22.047
```

```
alignment_block:
```

```
US-09-528-682-3 x US-09-477-574A-1 ..
```

```
Align seg 1/1 to: US-09-477-574A-1 from: 1 to: 3054
```

```
14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
||||| : : : : : |||||
892 ATTGAGATTCACACCCAGCCAGAGATAACCCAG..... 921
```

```
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly 47
||||| : : : : : |||||
922 .CCAATCACATCC.....ATCAAACTGGAGATAGTGACCTCCAAAAGTG 964
```

```
47 lYMetGlnAspGlyAsp.....LeuPhe 54
||| ||| |||||
965 GTGGCCAGCGGGGATCATGTCTCATGTCTGTCAGCAAGTGGGAGCCTAGCA 1014
```

```
55 GlyThrValAsnGlnSerAsnPhePro..... 63
```































; SEQ ID NO 79236  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Glycine max noir  
; FEATURE:  
; OTHER INFORMATION: Clone ID: uc-gmflLIB3275P095d05b1  
US-09-874-708A-79236

alignment\_scores:  
Quality: 76.00 Length: 99  
Ratio: 1.357 Gaps: 6  
Percent Similarity: 56.566 Percent Identity: 29.293

## alignment\_block:

US-09-528-682-3 x US-09-874-708A-79236/rev ..  
Align seg 1/1 to reverse of: US-09-874-708A-79236 from: 1 to: 600

```
14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :::::||||| |||||:::|||||::: :::: |||
551 CTGGAAGCCAACTAGAACACATCAACACGACCAAAAGACGTTGAGTG 502
   :::::||||| |||||:::|||||::: :::: |||
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
   ||| |||||:::|||||::: |||
501 CACGTTGCTTGGACGATATTCTACTTTTGTGGCGATCATAAGCCCTTGAT 452
   ||| ||||| ||||| ||||| ||||| ||||| |||||
43 ..SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
   ||| ||||| ||||| ||||| ||||| ||||| |||||
451 ATAGTTTTCGGCGAGGG.....GGAACCAATTTGGAATGGCAAC 411
   :::::||||| |||||:::|||||::: :::: |||
59 GlnSer...AsnPhePro.....MetSerTh 66
   :::::||||| |||||:::|||||::: :::: |||
410 ACTGCATTCATATATCCGTTCTTCATGTATATCTCCACACTATCAGG 361
   :::::||||| |||||:::|||||::: :::: |||
66 rPheGluGlnValProAsnAsn.....LysGluPheLysGlyValI 80
   :::::||||| |||||:::|||||::: :::: |||
360 GTATGAAGAACTCGATCTAAATCGAGTCCAGAAAGACGATTCGGGT... 315
   :::::||||| |||||:::|||||::: :::: |||
80 leSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArg 95
   :::::||||| |||||:::|||||::: :::: |||
314 .....GAAACTCGATGTGCTGCAGCTGAGCGGCGGATTCGA 276
   :::::||||| |||||:::|||||::: :::: |||
```

seq\_name: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:US-60-211-750-78103

seq\_documentation\_block:  
; Sequence 78103, Application US/60211750  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-211(51933)A  
; CURRENT APPLICATION NUMBER: US/60/211.750  
; CURRENT FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 90609  
; SEQ ID NO 78103  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Glycine max noir  
; OTHER INFORMATION: Clone ID: uc-gmflLIB3275P095d05b1  
US-60-211-750-78103

alignment\_scores:  
Quality: 76.00 Length: 99  
Ratio: 1.357 Gaps: 6  
Percent Similarity: 56.566 Percent Identity: 29.293

## alignment\_block:

US-09-528-682-3 x US-60-211-750-78103/rev ..

Align seg 1/1 to reverse of: US-60-211-750-78103 from: 1 to: 600

```
14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
   :::::||||| |||||:::|||||::: :::: |||
551 CTGGAAGCCAACTAGAACACATCAACACGACCAAAAGACGTTGAGTG 502
   :::::||||| |||||:::|||||::: :::: |||
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
   ||| |||||:::|||||::: |||
501 CACGTTGCTTGGACGATATTCTACTTTTGTGGCGATCATAAGCCCTTGAT 452
   ||| ||||| ||||| ||||| ||||| ||||| |||||
43 ..SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsn 58
   ||| ||||| ||||| ||||| ||||| ||||| |||||
451 ATAGTTTTCGGCGAGGG.....GGAACCAATTTGGAATGGCAAC 411
   :::::||||| |||||:::|||||::: :::: |||
59 GlnSer...AsnPhePro.....MetSerTh 66
   :::::||||| |||||:::|||||::: :::: |||
410 ACTGCATTCATATATCCGTTCTTCATGTATATCTCCACACTATCAGG 361
   :::::||||| |||||:::|||||::: :::: |||
66 rPheGluGlnValProAsnAsn.....LysGluPheLysGlyValI 80
   :::::||||| |||||:::|||||::: :::: |||
360 GTATGAAGAACTCGATCTAAATCGAGTCCAGAAAGACGATTCGGGT... 315
   :::::||||| |||||:::|||||::: :::: |||
80 leSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArg 95
   :::::||||| |||||:::|||||::: :::: |||
314 .....GAAACTCGATGTGCTGCAGCTGAGCGGCGGATTCGA 276
   :::::||||| |||||:::|||||::: :::: |||
```

seq\_name: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:US-08-484-688-2

seq\_documentation\_block:  
; Sequence 2, Application US/08484688  
; GENERAL INFORMATION:  
; APPLICANT: Pasloske, Brittan L.  
; APPLICANT: Baruch, Dror I.  
; APPLICANT: Howard, Russell J.  
; TITLE OF INVENTION: Malaria Vaccines  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Affymax Technologies, N.V.  
; STREET: 4001 Miranda Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,688  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/161,285  
; FILING DATE: 01-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kaster, Kevin R.  
; REGISTRATION NUMBER: 32,704  
; REFERENCE/DOCKET NUMBER: 1051  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-496-2300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1922 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-484-688-2

alignment\_scores:  
Quality: 76.00 Length: 93



```

Ratio: 1.462                                     Gaps: 3
Percent Similarity: 55.914      Percent Identity: 24.731

alignment_block:
US-09-528-682-3 x US-08-484-688-2      ..

Align seg 1/1   to: US-08-484-688-2 from: 1   to: 1922

      8  GlnglnAlaTyrGluProIleGluValasnThrAsnThrValThrGlnIl 24
          |||||  |||||  |||||  :||||:  :||||:  :||||:
1055  CAGTCAAAATAATGAAAAAATAAGGAAGAAGCTCGTAGCCAAATGAATG 1104
          |||||  |||||  |||||  :||||:  :||||:  :||||:
1105  TAACCAAAATAAATGACATTTCGTCTTATGAAAAAATAATTAAGAAGCTACT 1154
          |||||  |||||  |||||  :||||:  :||||:  :||||:
      40  laLeuIleSerAlaGluGlyGlyMetClnAspclyAspleupheGlyThr 56
          :||||  ::|  ::|  ::|  ||||
1155  CAATTAGATGATATACAAGAAAGAAAGTTGC..... 1186
          :||||  :||||  :||||  :||||  :||||  :||||
      57  ValAsnGlnSerAsnPhePrometSerThrpheGluInValProAsnAs 73
          :||||  :||||  :||||  :||||  :||||  :||||
1187  .....ATACCAGAAATATTATTTAAAAAATTACCATGAA 1221
          :||||  :||||  :||||  :||||  :||||  :||||
      73  nLysGluPheLysGlyValIleSerAlahsnVallYsrAspmetAsnP 90
          |  :||  :||  :||  :||  :||  :||
1222  TATAAATGAGATACATACATAATGAATCATAT.....AAAAATA 1259
          :||||  :||||  :||||  :||||  :||||  :||||
      90  heLysLysLeuLeuArgpheMetGluasp 99
          :||||  :||||  :||||  :||||  :||||  :||||
1260  ATAACAAATATTATAAATTTATGGAAGAT 1288
          :||||  :||||  :||||  :||||  :||||  :||||

seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-484-688A-2

seq_documentation_block:
; Sequence 2, Application US/08484688A
; GENERAL INFORMATION:
; APPLICANT: Pasloske, Brittan L.
; APPLICANT: Baruch, Dror I.
; APPLICANT: Howard, Russell J.
; TITLE OF INVENTION: Malaria Vaccines
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, N.V.
; STREET: 4001 Miranda Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484.688A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,285
; FILING DATE: 01-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1922 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: linear

```



```
US-09-528-682-3 x US-60-212-664-145
Align seg 1/1 to: US-60-212-664-145 from: 1 to: 181488
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94387 ATTAACACAAATTTGGTTTTCATTAATAAGGAGCAATATTTTCCCTCT 94436
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetG 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94437 GGAT.....TTTCTGTAGCCCAAGCACCACCAATTTGAAGAC 94471
49 InAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94472 TCGGTGAGCATATC.....ACTTCAGATGATAGCATGCTCCCAAGCAA 94515
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerA 82
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94516 GACAAATGAAGG.....GAAATGAAGCAAGGAGGAGTGGGGAG 94556
82 laAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPhe 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94557 CTACATAAAACCGAGCTATTCTTTGCCAAGTGCGCTATT 94600
seq_name: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-60-245-201-69
seq_documentation_block:
; Sequence 69, Application US/60245201
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000879
; CURRENT APPLICATION NUMBER: US/60/245,201
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 449058
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(449058)
; OTHER INFORMATION: n = A,T,C or G
US-60-245-201-69
alignment_scores:
Quality: 76.00 Length: 82
Ratio: 1.583 Gaps: 4
Percent Similarity: 58.537 Percent Identity: 35.366
alignment_block:
US-09-528-682-3 x US-60-245-201-69
Align seg 1/1 to: US-60-245-201-69 from: 1 to: 449058
16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67582 ATTAACACAAATTTGGTTTTCATTAATAAGGAGCAATATTTTCCCTCT 67631
32 uAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetG 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67632 GGAT.....TTTCTGTAGCCCAAGCACCACCAATTTGAAGAC 67666
49 InAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67667 TCGGTGAGCATATC.....ACTTCAGATGATAGCATGCTCCCAAGCAA 67710
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerA 82
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
67711 GACAAATGAAGG.....GAAATGAAGCAAGGAGAGTGGGGAG 67751
82 laAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPhe 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67752 CTACATAAAACCGAGCTATTCTTTGCCAAGTGCGCTGCTATT 67795
seq_name: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-60-127-461-2280
seq_documentation_block:
; Sequence 2280, Application US/60127461
; GENERAL INFORMATION:
; APPLICANT: Chapman, Rowan
; APPLICANT: Thornton, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF CANDIDA ALBICANS
; FILE REFERENCE: PM-0007-2 P
; CURRENT APPLICATION NUMBER: US/60/127,461
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 8247
; SOFTWARE: PERL Program
; SEQ ID NO 2280
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SKB807853286
US-60-127-461-2280
alignment_scores:
Quality: 75.00 Length: 88
Ratio: 1.562 Gaps: 2
Percent Similarity: 54.545 Percent Identity: 25.000
alignment_block:
US-09-528-682-3 x US-60-127-461-2280
Align seg 1/1 to: US-60-127-461-2280 from: 1 to: 636
14 IleGluValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluVa 30
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 TTGAATGTTGCTAAACCGCCGTCACAGAAA...TCTGGTCGAGACGATG 338
30 lProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyG 47
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
339 TATTATGCTGGGCTCTATTGGACCTTATGCTACTTTCCTGCCAACGGAT 388
47 lyMetGlnAspGlyAsp.....
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
389 CAGATATAACGCTGATTATCAAGGTGTTACTGACGAAGAGTTGATTGAA 438
53 .....LeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
439 TACCACACTCCGTTGTTGAAATTTTACGAGAAATTCGACCTTGACATAT 488
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerA 82
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
489 ATGCATTGAGACTATTCACAGCTTCCAAAGCTTTGAGGGTATATTTGGGG 538
82 laAsnValLysTyr 86
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
539 TGGCAAAAAGATAC 552
seq_name: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-614-150-42389
seq_documentation_block:
; Sequence 42389, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
```



alignment\_scores:

alignment\_block:

Align seq 1/1 to: US-60-167-324-1740 from: 1 to: 4528

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlyse 27  
:::||||| ||| ::|||  
2723 TACAATCCTCGTAGATTATGAAAAATCTAAGGAAGATGATAAACGGACA 2772  
27 rAsn.....G 29  
2773 AGACCCAAGTGTGCATCCGTGTCAGAAGACTTTTtaggcgcatgg 2822

```

29 luValProLeuAspGlyArgThr.....SerAsnPheAlaLeulle 42
|| |||||||
2823 AGATGTGTTTCGGATGCGCTTATATTCAGACGTGTGAACATTCAAATTTTG 2872
||||| : : : : : |||
43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGI 59
||||| : : : : : |||
2873 TCCGGA...AACCGTTTAGAAATCAGTGAACCTCCCTGTGGCGCTATGGAC 2919
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
||||| : : : : : |||
2920 GCAAAACTACAAGGAAAATGTCTCTGGAGCCTTTATCAACGGCACCGGAAA 2969
||||| : : : : : |||
76 .PheLysGlyValIleSerAlaAsnValLystyr.....AspMetAsn 89
||||| : : : : : |||
2970 AGGTTTAAGGTATATTATTCGAGTACAGGAGTATCATACAGACACCACC 3019
90.PheLysLysLeuLeuArgPheMetGluAspAspPheileGlyValHisGI 106
||||| : : : : : |||
3020 GTTCGCTTTGTGCATGTTTCGCCACTGGAGAATTTGAGCGCGATTCATGC 3069
106 yGlu 107
|||
3070 AGAA 3073

```

seq documentation block:

; GENERAL INFORMATION:

; TITLE OF INVENTION: Isolated Drosophila proteins, nucleic

FILE REFERENCE: CL000174

; CURRENT FILING DATE: 1999-12-28

```

; SOFTWARE: FastSEQ for Windows Version 4.0

```

seq\_documentation\_block:

# THE UNIVERSITY OF CHICAGO







```
27 rAsn.....G 29
   :
2773 AGAGCCAAGTGTGATCGTCGTTACAGAACTTTTAGGACGCATGG 2822
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
   |||
2823 AGTATGTTTCGGATGGTGTATATTCAGACTGGTAACATTCAAATTTTG 2872
43 SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
   |||
2873 TCCGGA...AACCGTTTGAAGAAATCGAGTACAGTCCCTGTGGCGCATGGAC 2919
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
   :
2920 GCAAAACTACAAGGAAATGCTCGGAGCCCTTATCAACAGCCGCCGAAA 2969
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
   |||
2970 AGGTTAAGGGTATTATTCCGAGTACAGGAGTATCATACAGACACCACC 3019
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
   :
3020 GTTCGCTTTGTGATCAGTTTCGCACCTGGAGAAATTTGAGCGCATTCATGC 3069
106 yGlu 107
   :
3070 AGAA 3073
```

seq\_name: /cgn2\_6/ptodata/2/pna/us6019\_COMB.seq:us-60-191-637-42026

```
seq_documentation_block:
; Sequence 42026, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42026
; LENGTH: 4528
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-637-42026
```

```
alignment_scores:
  Quality: 75.00      Length: 118
  Ratio: 1.230       Gaps: 5
  Percent Similarity: 51.695      Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x US-60-191-637-42026

Align seg 1/1 to: US-60-191-637-42026 from: 1 to: 4528

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlyse 27
   |||
2723 TACAATCCTCGTGAGATTATGAAAAATCTAAGGAAGATGATAAACGGACA 2772
27 rAsn.....G 29
   :
2773 AGAGCCAAGTGTGATCGTCGTTACAGAACTTTTAGGACGCATGG 2822
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
   |||
2823 AGTATGTTTCGGATGGTGTATATTCAGACTGGTAACATTCAAATTTTG 2872
43 SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
   |||
2873 TCCGGA...AACCGTTTGAAGAAATCGAGTACAGTCCCTGTGGCGCATGGAC 2919
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
   :
2920 GCAAAACTACAAGGAAATGCTCGGAGCCCTTATCAACAGCCGCCGAAA 2969
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
   |||
2970 AGGTTAAGGGTATTATTTCGAGTACAGGAGTATCATACAGACACCACC 3019
```

```
2873 TCCGGA...AACCGTTTGAAGAAATCAGTGAATCCCTGTGGCGCATGGAC 2919
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
   :
2920 GCAAAACTACAAGGAAATGCTCGGAGCCCTTATCAACAGCCGCCGAAA 2969
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
   |||
2970 AGGTTAAGGGTATTATTCCGAGTACAGGAGTATCATACAGACACCACC 3019
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
   :
3020 GTTCGCTTTGTGATCAGTTTCGCACCTGGAGAAATTTGAGCGCATTCATGC 3069
106 yGlu 107
   :
3070 AGAA 3073
```

seq\_name: /cgn2\_6/ptodata/2/pna/us6019\_COMB.seq:us-60-191-700-1769

```
seq_documentation_block:
; Sequence 1769, Application US/60191700
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; FILE REFERENCE: CL000391
; CURRENT APPLICATION NUMBER: US/60/191,700
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 2211
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1769
; LENGTH: 4528
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-700-1769
```

```
alignment_scores:
  Quality: 75.00      Length: 118
  Ratio: 1.230       Gaps: 5
  Percent Similarity: 51.695      Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x US-60-191-700-1769

Align seg 1/1 to: US-60-191-700-1769 from: 1 to: 4528

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlyse 27
   |||
2723 TACAATCCTCGTGAGATTATGAAAAATCTAAGGAAGATGATAAACGGACA 2772
27 rAsn.....G 29
   :
2773 AGAGCCAAGTGTGATCGTCGTTACAGAACTTTTAGGACGCATGG 2822
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
   |||
2823 AGTATGTTTCGGATGGTGTATATTCAGACTGGTAACATTCAAATTTTG 2872
43 SerAlaGluGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
   |||
2873 TCCGGA...AACCGTTTGAAGAAATCAGTGAATCCCTGTGGCGCATGGAC 2919
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
   :
2920 GCAAAACTACAAGGAAATGCTCGGAGCCCTTATCAACAGCCGCCGAAA 2969
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
   |||
2970 AGGTTAAGGGTATTATTTCGAGTACAGGAGTATCATACAGACACCACC 3019
```







```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 7767
; TYPE: DNA
; ORGANISM: Drosophila
US-60-175-871-1811

```

alignment_scores:		
Quality:	75.00	Length: 118
Ratio:	1.230	Gaps: 5
Percent Similarity:	51.695	Percent Identity: 25.424

```

alignment_block:
US-09-528-682-3 x US-60-175-871-1811 ..
Align seg 1/1 to: US-60-175-871-1811 from: 1 to: 7767
11 TyrGluPtoIleGluValAsnThrValThrGlnIleAsnGlySe 27
||||| ||| :|: ||||| ||| :|:
4881 TACATCTCGAGATTGAAAAATCTAAGGAAGATGATAACGGACA 4930
27 rAsn.....G 29
:|:
4931 ACAGGCCAACTGTGATGCATCCGTGGTACAAAGACTTTTAGGACGCATGG 4980

```

seq\_name: /cqn2\_6/ptodata/2/pna/us6018\_COMB.seq:US-60-184-775-1653

```

seq_documentation_block:
; Sequence 1653, Application US/60184775
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000282
; CURRENT APPLICATION NUMBER: US/60/184,775
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 2083
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1653
; LENGTH: 7767
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-184-775-1653

```

```
alignment_scores:
  Quality: 75.00
  Length: 118
```

```

Ratio: 1.230      Gaps: 5
Percent Similarity: 51.695      Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x US-60-184-775-1653      ..

Align seg 1/1 to: US-60-184-775-1653 from: 1 to: 7767

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlySe 27
   |||::|||   |||::   |||   ::||||||::
4881 TACAATCCTCGTAGATTATGAAAAATCTAAGGAAGATGATAAACGGACA 4930

27 rAsn.....G 29
   ::::
4931 AGAGCCAAAGTGTGATGCATCCGTGTACAGAAGCTTTTAGCAGCGATGG 4980

29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
   || ||| ||| ||| ||| ||| ::|||:::
4981 AGTATGTTTCGATGGTCTTATATTCACACGTGGTAACATTCAAATTTTG 5030

43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGl 59
   |||::: |||::: |||::: |||
5031 TCCGGA...AACCGTTTAGAAATCAGTGAACCTCCCTGTGGCGGTATGGAC 5077

59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
   ::|||::: ::|||::: |||::: |||:::
5078 GCAAACTACAAAGGAAAATGTCTGGAGCCTTTATCAAACGGCAGCAGAA 5127

76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
   ||| ||| ::||| ||| ||| ::||| ::|
5128 AGGTTAAGGGTATTATTTCGAGTACAGGAGATTCATACAGACACACC 5177

90 PheLysLysLeuAurGppPheMetGluAspAspPheIleGlyValHISgl 106
   ::::: |||::: |||::: |||:::
5178 GTTCGCTTTGTGATCAGITTCGCACCTGGAGAAATTTGAGCGCATTCATGC 5227

106 yGlu 107
   :|||
5228 AGAA 5231

```

```

seq_name: /cqn2_5/ptoddata/2/pna/US096A_COMB.seq:US-09-614-150-42388

seq_documentation_block:
; Sequence 42388, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42388
; LENGTH: 7769

```







```

2837 AGAGCCAAAGTGTGATCGATCGGTACAGAACTTTTATAGAGCGCATGG 2788
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
2787 AGTATGTTCCGATGGTGGTGTATATTCAGACTGGTAACATTCAAAATTTG 2738
43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
2737 TCCGGA...AACCGTTTAGAAATCAGTGAACCTCCTGTGGCGGTATGGAC 2691
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
2690 GCAAAACTACAGGAAATGCTCGGAGCCTTTATCAACGGCACCGCGAAA 2641
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
2640 AGGTTAAGGGTATTATTTCCGAGTACAGGAGTATCATACAGACACCCACC 2591
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
2590 GTTCGCTTTGTCATCAGTTTCGCACCTGGAGAATTTGAGCGCATTCATGC 2541
106 yglu 107
2540 AGAA 2537
seq_name: /cqn2_6/ptodata/2/pna/US6015_COMB.seq:US-60-150-584-497

```

```

seq_documentation_block:
; Sequence 497, Application US/60150584
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CL000087
; CURRENT APPLICATION NUMBER: US/60/150,584
; CURRENT FILING DATE: 1999-08-25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 497
; LENGTH: 13844
; TYPE: DNA
; ORGANISM: Drosophila
US-60-150-584-497

```

```

alignment_scores:
  Quality: 75.00      Length: 118
  Ratio: 1.230       Gaps: 5
  Percent Similarity: 51.695  Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x US-60-150-584-497

Align seg 1/1 to: US-60-150-584-497 from: 1 to: 13844

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlySe 27
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5151 TACAATCCTCGTGAATATGAAAAATCTAAGGAAGATGATAACGGACA 5200
27 rAsn.....G 29
5201 AGAGCCAAAGTGTGATCGATCGGTACAGAACTTTTATAGAGCGCATGG 5250
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5251 AGTATGTTCCGATGGTGGTGTATATTCAGACTGGTAACATTCAAAATTTG 5300
43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
5301 TCCGGA...AACCGTTTAGAAATCAGTGAACCTCCTGTGGCGGTATGGAC 5347
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
5348 GCAAAACTACAGGAAATGCTCGGAGCCTTTTATCAACGGCACCGCGAAA 5397
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
5398 AGGTTAAGGGTATTATTTCCGAGTACAGGAGTATCATACAGACACCCACC 5447
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
5448 GTTCGCTTTGTCATCAGTTTCGCACCTGGAGAATTTGAGCGCATTCATGC 5497

```

```

59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
5348 GCAAAACTACAGGAAATGCTCGGAGCCTTTTATCAACGGCACCGCGAAA 5397
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
5398 AGGTTAAGGGTATTATTTCCGAGTACAGGAGTATCATACAGACACCCACC 5447
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
5448 GTTCGCTTTGTCATCAGTTTCGCACCTGGAGAATTTGAGCGCATTCATGC 5497
106 yglu 107
5498 AGAA 5501
seq_name: /cqn2_6/ptodata/2/pna/US6015_COMB.seq:US-60-150-584-508

```

```

seq_documentation_block:
; Sequence 508, Application US/60150584
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CL000087
; CURRENT APPLICATION NUMBER: US/60/150,584
; CURRENT FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 1072
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 508
; LENGTH: 13844
; TYPE: DNA
; ORGANISM: Drosophila
US-60-150-584-508

```

```

alignment_scores:
  Quality: 75.00      Length: 118
  Ratio: 1.230       Gaps: 5
  Percent Similarity: 51.695  Percent Identity: 25.424

alignment_block:
US-09-528-682-3 x US-60-150-584-508

Align seg 1/1 to: US-60-150-584-508 from: 1 to: 13844

11 TyrGluProIleGluValAsnThrAsnThrValThrGlnIleAsnGlySe 27
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5151 TACAATCCTCGTGAATATGAAAAATCTAAGGAAGATGATAACGGACA 5200
27 rAsn.....G 29
5201 AGAGCCAAAGTGTGATCGATCGGTACAGAACTTTTATAGAGCGCATGG 5250
29 luValProLeuAspGlyArgTyr.....SerAsnPheAlaLeuIle 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5251 AGTATGTTCCGATGGTGGTGTATATTCAGACTGGTAACATTCAAAATTTG 5300
43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
5301 TCCGGA...AACCGTTTAGAAATCAGTGAACCTCCTGTGGCGGTATGGAC 5347
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGlu. 75
5348 GCAAAACTACAGGAAATGCTCGGAGCCTTTTATCAACGGCACCGCGAAA 5397
76 ..PheLysGlyValIleSerAlaAsnValLysTyr.....AspMetAsn 89
5398 AGGTTAAGGGTATTATTTCCGAGTACAGGAGTATCATACAGACACCCACC 5447
90 PheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHisG1 106
5448 GTTCGCTTTGTCATCAGTTTCGCACCTGGAGAATTTGAGCGCATTCATGC 5497

```







```

Ratio: 1.307      Gaps: 3
Percent Similarity: 55.882      Percent Identity: 24.510

alignment_block:
US-09-528-682-3 x US-09-815-264-92808      ..

Align seg 1/1 to: US-09-815-264-92808 from: 1 to: 795

13  ProfileValAsnThrAsnThrValThrGlnIleAsnGlySerAsnG1 295
||||| :|||:|||||:||||| :|||:||||| :|||:|||||
388  CCAATTGTTCTTCACACACATATCTCAATTGCAGGTCAATTGCAGCCAGAA 437
29  uValProLeuAspGlyArgTyrSerAsnPheAlaLeuIle..... 42
:|||| :|||:|||||:||||| :|||:||||| :|||:|||||
438  AGAGCGGATGGATTTCAAAGTCGGTTCCTTCTCCTACTCCTCTTAATCG 487
43  .....SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGly 55
:||||:||||:||||:||||:||||:||||:||||:||||:
488  TCACCTGCGGAGCTGCCTCAAGGT.....CAAGGCAATGATTTTGGT 528
56  ThrValAsnGlnSerAsnPheProMetSerThrPheGluGlnValProAs 72
||| :|||:|||||:||||| :||| :|||
529  ACGGCCTTAATCCACACACTTCCTCTCCTCCTACCTATGCTCCTCCT.. 576
72  nAsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAspMetA 89
:||||:||||:||||:||||:||||:||||:||||:||||:
577  .....GGCTCCTGCATGCAGTTTGGGAACCA 601
89  snPheLysLysLeuLeuArgPheMetGluAspPheIleGlyValHis 105
|||||:||||:||||:||||:||||:||||:||||:||||:
602  ACATGTCAACATTTCTCAAGGGAAATTTGGACAAGTTTTTTATTGTGCGGT 651

106 GlyGlu 107
:||||
652  AGTAA 657

```











OM of: US-09-528-682-3 to: Pending\_Patents\_NA\_New: \* out\_format : pfs

Date: Jun 18, 2002 9:49 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09528682/runat\_18062002\_082444\_7900/app\_query.fasta\_1.689  
-DB=Pending\_Patents\_NA\_New -QMT=fastcap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000  
-LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMF=pts -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682@cgn1\_1\_572 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT\_THREADS=1

Search information block:

Query: US-09-528-682-3

Query length: 107

Database: Pending\_Patents\_NA\_New: \*

Database sequences: 1014543

Database length: 72792371

Search time (sec): 403.930000

score\_list:

Sequence	Strd Orig	zScore	EScore	Len	! Documentation
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-168431 -	78.50	171.77	0.1871		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3396 +	78.00	175.54	0.1154		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-663-203708 -	76.50	172.46	0.1713	11	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-053-3758A-1300 +	72.50	143.93	6.65		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-27154 +	72.00	154.36	1.75		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25440 +	71.50	152.36	2.25		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-284619 +	70.50	163.21	0.5609		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-286485 +	70.50	163.07	0.5706		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203709 -	70.50	162.94	0.5803		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-285486 +	70.50	161.67	0.6828		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25495 +	70.00	141.40	9.19		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-6689 -	69.00	167.13	0.3390		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-34079 -	69.00	156.46	1.33		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-34080 -	69.00	156.46	1.33		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-699-2010 +	68.50	161.83	0.6691		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-32947 +	67.50	152.02	2.36		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-12794 -	67.50	144.40	6.26		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25839 +	66.50	139.52	11.71		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-380-1 +	66.00	132.13	30.19	428	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-103464 +	65.50	148.40	3.75		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4162 +	65.50	147.86	4.02		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25694 +	65.50	139.45	11.80		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-45758 +	65.50	139.45	11.80		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-45811 +	65.50	139.45	11.80		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3205 +	65.50	137.23	15.69		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-464 +	65.50	133.99	23.77	3	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-42147 +	65.50	129.43	42.66		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-6867 +	65.00	145.46	5.46		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-24073 +	64.50	141.21	9.42		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-425 -	64.50	139.73	11.39		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-882-227-521 +	64.50	134.06	23.57		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-3313-3135 -	64.50	131.31	121.00	8	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-930-213-248 +	64.00	152.78	2.14		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-131-129-15 +	63.50	139.27	12.08	5	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1978 +	63.50	132.67	28.17		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-953-1 +	63.50	129.20	43.95	298	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-1227 +	63.50	128.02	51.16		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-634-754C-766 -	63.00	144.87	5.90		

/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-168431 -	63.00	140.36	10.51		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-121 -	63.00	134.06	23.57		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-330 +	63.00	132.79	27.73		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-375B-832 +	63.00	117.29	202.43		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-13583 +	62.50	149.04	3.45		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-34787 +	62.50	137.68	14.82		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-46330 +	62.50	134.03	23.67		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-33321 +	62.50	132.82	27.65		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-36604 +	62.50	131.08	34.56		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-045-815-5 +	62.50	128.68	48.98		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-152056 +	62.00	136.83	16.52		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-113202 -	62.00	135.24	73.07		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-601 -	62.00	121.50	118.07		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-5060 +	61.50	130.79	35.87		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-045-815-3 +	61.50	125.39	71.68		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-045-815-1 +	61.50	124.57	79.65		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-103402 +	61.00	125.18	73.63		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-103-76 +	61.00	124.68	78.52		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41472 +	61.00	123.87	87.06		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-153 +	61.00	123.86	87.23		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-32945 +	60.50	132.07	30.43		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-33132 +	60.50	132.07	30.43		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-34610 +	60.50	131.97	30.80		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41273 +	60.50	131.78	31.56		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-45837 +	60.50	127.22	56.66		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40326 +	60.50	126.69	60.67		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-45847 +	60.50	119.39	154.72		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-24843 +	60.50	116.19	233.13		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-3310-5544 +	60.00	142.99	7.50		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-24123 -	60.00	132.29	29.59		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-12037 +	60.00	131.56	32.50		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37269 +	60.00	126.13	65.19		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3381 +	60.00	125.20	73.44		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-340-209B-3840 +	60.00	121.02	125.49		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-174961 -	60.00	52.23	6.9e+		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-668 -	59.50	136.94	16.30		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-11904 -	59.50	135.62	19.31		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-059-964-49 +	59.50	130.53	37.06		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-6772 -	59.50	130.06	39.38		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-732 -	59.50	129.86	40.42		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-11979 -	59.50	129.86	40.42		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-346 -	59.50	129.83	40.57		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-11224 -	59.50	129.83	40.57		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3653 +	59.50	129.67	41.41		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-14837 -	59.50	127.74	53.03		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-24650 +	59.50	127.17	57.03		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40139 +	59.50	125.05	70.62		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-15161 +	59.50	125.05	74.91		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-14877 -	59.50	124.31	82.30		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-882-227-619 +	59.50	121.30	121.07		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-340-209B-4777 +	59.50	120.46	134.80		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-108-605-234 +	59.50	115.09	268.48		
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12774 -	59.50	113.34	336.15		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-1467 +	59.50	112.50	374.56		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40091 +	59.00	127.61	53.92		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37172 +	59.00	127.08	57.73		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-4581 +	59.00	126.42	61.60		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40722 +	59.00	126.42	62.77		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-10206 +	59.00	125.88	67.28		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-22667 +	59.00	125.88	67.28		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-26028 +	59.00	125.43	71.26		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-340-209B-234 +	59.00	124.76	77.73		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1297 +	59.00	121.86	112.66		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-895-913A-9 +	59.00	119.78	147.24		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1168 +	59.00	118.98	163.05		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-472 -	59.00	117.99	185.06		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-22879 +	58.50	139.66	11.20		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-991-936-970 -	58.50	137.47	15.22		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40892 +	58.50	129.38	42.96		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-45059 +	58.50	123.82	87.64		
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-27018 +	58.50	121.61	116.42		
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-053-375B-1208 +	58.50	118.79	167.20		



[illegible]



/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-5291 - 56.50 107.00 758.08  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-135-985-3 + 56.50 96.00 3.1e+03 141  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-147-196-3 + 56.50 96.00 3.1e+03 141  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US01-45053-13 - 56.50 80.30 2.3e+04 70  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-135-322-13 - 56.50 80.30 2.3e+04 70  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-174581 - 56.50 57.79 3.8e+05  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-141-530-1 + 56.00 128.31 49.27 449  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-1514 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-1564 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-6514 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-6564 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-1514 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-1564 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-6514 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-6564 - 56.00 125.15 73.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-281076 + 56.00 125.01 75.23  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:US-10-027-632-11152 + 56.00 123.37 92.92  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-2517 - 56.00 121.19 122.89  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-2517 - 56.00 121.19 122.89  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:US-10-027-632-119672 - 56.00 119.64 149.77  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-540-209B-4393 + 56.00 118.59 171.47  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-540-209B-1202 + 56.00 118.26 178.89  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-935-625-362 + 56.00 117.46 198.25  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-935-625-3076 + 56.00 117.46 198.25  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-935-625-11735 + 56.00 117.46 198.25  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-935-625-16403 + 56.00 117.46 198.25  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-935-625-28123 + 56.00 117.46 198.25  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-372-508-7 - 56.00 116.27 230.99 154  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-60-360-039-41098 + 56.00 116.03 238.11  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-919-002-13017 + 56.00 115.36 259.52  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-053-375B-627 + 56.00 115.18 265.65  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-105-698-363 + 56.00 115.03 270.65  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-053-375B-481 - 56.00 114.36 294.75  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-442-384B-484 - 56.00 114.36 294.75  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-919-002-6574 - 56.00 114.22 300.25  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-314 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-364 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-5514 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-5564 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-514 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-5514 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-5564 - 56.00 111.12 446.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-007-926A-295 + 56.00 109.95 519.45  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-1295 + 56.00 109.85 526.26  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-2295 + 56.00 109.85 526.26  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-1295 + 56.00 109.85 526.26  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-2295 + 56.00 109.85 526.26  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-60-360-039-45594 - 56.00 108.94 591.07  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-102-524-1854 + 56.00 108.44 630.52  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-102-524-1854 + 56.00 108.44 630.52  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-7295 + 56.00 108.01 665.62  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-102-524-1853 + 56.00 105.94 868.36  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-7295 + 56.00 105.94 868.36  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-102-524-1852 + 56.00 105.73 890.26  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-6295 + 56.00 105.66 900.34  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-6295 + 56.00 105.66 900.34  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-295 + 56.00 104.85 998.39  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-295 + 56.00 104.85 998.39  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:US-09-919-002-6552 - 56.00 104.50 1.0e+03  
/cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-5295 + 56.00 102.14 1.4e+03  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-5295 + 56.00 102.14 1.4e+03  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-105-299-12937 - 56.00 88.58 8.0e+03  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-704-302A-1055 + 55.50 129.45 42.58  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-704-302A-1058 + 55.50 127.29 58.46  
/cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-704-302A-1063 + 55.50 126.98 58.46  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-280956 - 55.50 123.72 88.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-280957 - 55.50 123.72 88.77  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-280958 - 55.50 123.61 90.00  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-253639 - 55.50 123.18 95.15  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-253639 - 55.50 122.12 107.81  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-20610 + 55.50 121.75 114.25  
/cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-20611 + 55.50 121.75 114.25











/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-35298-	53.00	106.25	834.64	52.50	117.04	209.0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-442-384B-747+	53.00	105.28	945.52	52.50	117.01	209.9
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-119-635-7+	53.00	104.91	990.63	52.50	117.01	209.9
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-47146+	53.00	104.58	1.0e+03	52.50	116.94	211.82
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-838-564A-1-	53.00	104.20	1.1e+03	52.50	116.62	220.66
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-1538-	53.00	104.12	1.1e+03	52.50	116.47	224.8
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-506+	53.00	103.94	1.1e+03	52.50	116.43	226.2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4577+	53.00	103.77	1.1e+03	52.50	116.17	233.82
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-46194+	53.00	103.51	1.2e+03	52.50	115.83	244.2
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-40359+	53.00	103.24	1.2e+03	52.50	115.83	244.2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-8236+	53.00	102.98	1.3e+03	52.50	115.81	244.75
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8163-	53.00	102.98	1.3e+03	52.50	115.81	244.75
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8174-	53.00	102.17	1.4e+03	52.50	115.46	256.25
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25513+	53.00	102.07	1.4e+03	52.50	115.46	256.25
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-3012+	53.00	101.67	1.5e+03	52.50	114.32	296.33
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-465+	53.00	101.21	1.6e+03	52.50	114.32	296.33
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-12012+	53.00	100.93	1.7e+03	52.50	114.02	308.27
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-2951-	53.00	99.16	2.1e+03	52.50	113.98	309.7
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-757-	53.00	99.16	2.1e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8173-	53.00	99.15	2.1e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8172-	53.00	98.17	2.4e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-09-909-567B-7+	53.00	98.17	2.4e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-089-647-56+	53.00	98.17	2.4e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-35176+	53.00	96.15	3.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8161+	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8162-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8172-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8173-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8174-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8175-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8176-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8177-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8178-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8179-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8180-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8181-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8182-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8183-	53.00	90.81	6.0e+03	52.50	113.90	312.78
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8184-	53.00	89.66	7.0e+03	52.50	113.83	315.79
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8836+	53.00	88.25	8.4e+03	52.50	113.80	316.79
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8163-	53.00	88.25	8.4e+03	52.50	113.39	333.96
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8175-	53.00	88.25	8.4e+03	52.50	113.39	333.96
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8176-	53.00	88.25	8.4e+03	52.50	113.39	333.96
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-8177-	53.00	79.22	2.7e+04	52.50	113.16	344.1
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12035-	53.00	73.91	5.2e+04	52.50	112.86	357.48
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12011+	53.00	63.27	2.0e+05	52.50	112.57	370.91
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-21686	52.50	124.95	75.84	52.50	112.49	375.06
/cgn2_6/ptodata/1/pna						



```
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2818 + 52.50 100.15 1.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-112613 + 52.50 99.89 1.9e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4118 + 52.50 99.26 2.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-143-155-1 + 52.50 99.23 2.1e+03 351
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-814 + 52.50 97.63 2.5e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-053-375B-702 + 52.50 97.27 2.6e+03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-25405 + 52.50 97.09 2.7e+03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-26966 + 52.50 94.20 3.9e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-125-818-1 + 52.50 90.80 8.1e+03 835
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-110-610-1 + 52.50 88.43 8.2e+03 106
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-9636 + 52.50 84.43 1.4e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-9637 + 52.50 84.43 1.4e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-14828 + 52.50 84.40 1.4e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-9635 + 52.50 81.55 2.0e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-003-806-10 + 52.50 61.23 2.5e+05 17
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331-2415 + 52.00 126.48 62.36
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-806B-3923 + 52.00 123.43 92.16
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-24472 + 52.00 121.92 111.82
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-11626 + 52.00 120.34 137.03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-43704-2553 + 52.00 118.67 169.76
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-154-2314 + 52.00 118.44 174.84
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-43450 + 52.00 117.32 201.71
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-6065 + 52.00 116.64 220.21
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-92141 + 52.00 116.62 220.75
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-304670 + 52.00 116.62 220.75
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-097-105-491 + 52.00 115.74 247.19 5
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-273197 + 52.00 115.74 247.19
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-180298 + 52.00 115.63 250.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-2959 + 52.00 114.48 290.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-116-355-27 + 52.00 114.34 295.55 65
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4497 + 52.00 113.11 346.13
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-34703 + 52.00 113.07 347.89
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-142859 + 52.00 112.35 381.51
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-156657 + 52.00 112.08 395.24
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-156658 + 52.00 112.08 395.24
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-376659 + 52.00 112.08 395.24
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-37178 + 52.00 112.00 399.44
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-37102 + 52.00 111.90 404.25
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-6671 + 52.00 111.51 424.83
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-09-882-232-160293 + 52.00 111.51 424.83
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-882-232-27419 + 52.00 111.43 429.70 8
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-27129 + 52.00 110.57 479.62
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-26567 + 52.00 110.08 510.34
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-42521 + 52.00 109.86 525.53
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-250762 + 52.00 109.66 538.88
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-36400 + 52.00 109.60 543.35
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-1225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-2225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-6225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-7225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-1225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-2225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-6225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-7225 + 52.00 109.52 548.47
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-255452 + 52.00 109.50 549.75
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-255453 + 52.00 109.50 549.75
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-13240-9 + 52.00 109.00 586.48 113
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-211 + 52.00 108.31 640.80
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-27436 + 52.00 108.31 640.80
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-256923 + 52.00 108.27 644.10
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3216 + 52.00 107.99 667.97
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-442-384B-430 + 52.00 107.70 692.69
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-12382 + 52.00 107.69 694.03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-41540 + 52.00 107.01 756.97
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-588-4 + 52.00 106.89 768.59 140
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-012-269-4 + 52.00 106.89 768.59 140
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-44764 + 52.00 106.86 771.32
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3932 + 52.00 106.84 773.38
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-574 + 52.00 106.62 796.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-5088 + 52.00 106.62 796.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-19667 + 52.00 106.62 796.06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-28139 + 52.00 106.62 796.06
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-06-360-039-25795 + 52.00 106.52 806.42
```

```
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-27533 - 52.00 106.25 834.16
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-003-698-3410 + 52.00 105.77 888.08
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-625-5153 + 52.00 105.68 897.96
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2937 + 52.00 105.49 919.90
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-13081 + 52.00 105.33 939.10
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-15128 + 52.00 105.29 943.38
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-253418 + 52.00 105.16 959.8
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-253140 + 52.00 105.15 960.5
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-15731 + 52.00 105.14 961.95
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-9758 + 52.00 105.02 977.01
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-103-239-1546 + 52.00 104.96 984.92
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-33458 - 52.00 104.74 1.0e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-40164 + 52.00 104.73 1.0e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4931 + 52.00 104.63 1.0e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-6788 + 52.00 104.07 1.1e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1749 + 52.00 104.02 1.1e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-27210 - 52.00 103.87 1.1e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-23796 + 52.00 103.00 1.3e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-4480 + 52.00 102.65 1.3e+0
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-108-580-1 - 52.00 102.65 1.3e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-3714 + 52.00 102.46 1.4e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-1402 + 52.00 102.01 1.4e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-256922 + 52.00 101.52 1.5e+
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-053-375B-1259 + 52.00 101.15 1.6e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-42055 + 52.00 101.11 1.6e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-24365 + 52.00 101.09 1.6e+0
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-1631 + 52.00 101.08 1.6e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-21010 + 52.00 101.02 1.6e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-25861 + 52.00 101.02 1.6e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-25059 + 52.00 101.01 1.6e+0
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-145-541-1 + 52.00 100.98 1.6e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-588-20 + 52.00 100.95 1.6e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-012-269-20 + 52.00 100.95 1.6e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-895-913A-197 + 52.00 100.26 1.8e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-2225 + 52.00 99.26 2.0e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-5225 + 52.00 99.26 2.0e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-2225 + 52.00 99.26 2.0e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-5225 + 52.00 99.26 2.0e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-630-630B-13 + 52.00 98.44 2.3e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-116-802-197 - 52.00 98.22 2.3e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-123-692-7 + 52.00 98.18 2.3e+03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-46003 + 52.00 98.14 2.4e+03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-09-053-375B-466 + 52.00 97.94 2.4e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-053-375B-377 - 52.00 97.58 2.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-331 - 52.00 97.57 2.5e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-103-313-240 - 52.00 97.41 2.6e+03
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-25331 + 52.00 96.92 2.8e+0
/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:US-60-360-039-24908 + 52.00 96.71 2.8e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2309 - 52.00 96.66 2.9e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-4433 + 52.00 96.66 2.9e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-909-567B-5 + 52.00 95.87 3.2e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-1839 - 52.00 95.30 3.4e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-548-409B-10 + 52.00 94.84 3.6e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-103-310-20 + 52.00 94.51 3.8e+0
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-103-310-20 + 52.00 94.41 3.8e+03
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-11253 + 52.00 94.32 3.9e+0
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-503-477A-17 + 52.00 93.32 4.4e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-916-963-39 + 52.00 90.75 6.1e+03
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-3489 + 52.00 90.62 6.2e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-13240-8 + 52.00 88.03 8.6e+03
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:US-10-105-299-12084 - 52.00 86.71 1.0e+0
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-13240-15 + 52.00 86.71 1.0e+0
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-310-18 - 52.00 76.12 3.9e+04
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-351-1 + 52.00 67.74 1.1e+05
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-1 - 52.00 60.59 2.7e+05
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-345-882-1 - 52.00 60.59 2.7e+05
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-126-704-1 - 52.00 60.59 2.7e+05
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-939-964A-1 - 52.00 48.96 9.6e+05
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-811-352B-1 - 52.00 41.18 1.7e+06
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-924-1 - 52.00 41.18 1.7e+06
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-067-514-1 - 52.00 37.77 1.9e+06
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-11160 - 51.50 121.44 118.8
```



```
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-12834 + 51.50 119.44 153.71
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-274596 + 51.50 118.51 173.21
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-274597 + 51.50 118.51 173.21
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-274598 + 51.50 118.51 173.21
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-22825 + 51.50 117.76 190.79
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-21966 + 51.50 117.26 203.29
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-11024 + 51.50 117.01 209.90
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-7958 + 51.50 116.36 228.14
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:PCT-US01-43704-1638 + 51.50 116.13 234.91
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-704-302A-1060 + 51.50 116.01 238.61
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-43664 - 51.50 115.56 252.93

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3396

seq_documentation_block:
; Sequence 3396, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3396
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-3396

alignment_scores:
Quality: 78.50 Length: 100
Ratio: 1.481 Gaps: 6
Percent Similarity: 53.000 Percent Identity: 29.000

alignment_block:
US-09-528-682-3 x US-09-540-209B-3396 ..
Align seg 1/1 to: US-09-540-209B-3396 from: 1 to: 2007
8 GlnGlnAlaTyrGluProIleGluValAsnThrValThrGlnII 24
550 CAAAAATCATATCGT...CTGAAATCCCGACTAATGGTATGACGGACAT 596
24 eAsnGlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAlaL 41
597 GAATGGTGGCTCGAGTGGCCCA...GATGATACTAT..... 630
41 euIleSerAlaGluGlyMetGlnAspGlyAspLeu.....Phe 54
631 .....GCTGAAGTGGGAAGTAACGAAGCAAACTGATGCTTACTTT 672
55 GlyThrValAsnGlnSerAsnPheProMetSerThrPheGluGlnValPr 71
673 GGTGCTGTAATAATCTCTTTTGGAT..... 699
71 oAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAspM 88
700 .....GCTATTATTATAGAGCCCAATATCGTGCTGATG 733
88 etAsnPheLysLysLeuLeuArgPheMetGluAspAspPheIleGlyVal 104
734 CTTCTTTT.....CGCTTTCATAAGGATAATCGTTGGGTGTT 771

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203708

seq_documentation_block:
; Sequence 203708, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
```

```
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 203708
LENGTH: 1194
TYPE: DNA
ORGANISM: Human
US-10-027-632-203708

alignment_scores:
Quality: 78.00 Length: 113
Ratio: 1.300 Gaps: 5
Percent Similarity: 53.097 Percent Identity: 25.664

alignment_block:
US-09-528-682-3 x US-10-027-632-203708/rev ..
Align seg 1/1 to reverse of: US-10-027-632-203708 from: 1 to: 1194
2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
863 TTCTTCAGCCAGCATCCACAGCCACAGCAGCTCTGAGTATTTCCT 814
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
813 TTCTTCAGTCACCCACATAAATGGCTTCAGGCGCTTCTGGGAAGCCT. 765
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMet 48
764 ..GAAGGAAGATTACAGCATACACTTGTGGYAGCATTAAGAGC..... 723
49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
722 .....TTCACCTCTTCC 712
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerA 82
711 TCAAGGATCCAATCTCCCTCAGTCAGTCAAGAAGCTCCAGGTATCTGAAC 662
82 la.....AsnValLysTyrAspMetAsnPheLysLysLeu.... 93
661 GATGCCAGGTCAATAAATCCCACTATGGTCACTCCATCAGGTCTCTGTCC 612
94 .....LeuArgPheMetGluAspAspPheIleGlyVal 104
611 TCAGAACTAGAGTAGTAAAGATAGACTCATGGAGTC 573

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-663-870A-5

seq_documentation_block:
; Sequence 5, Application US/09663870A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
```







```
3 PheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAs 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
928 TTCAACGAGTCTCTAGCAACGCGTCAAGGAGTTT...GTGAACCGGAA 974

19 nThrValThrGlnIleAsnGlySerAsnGluValPro...LeuAspGlyA 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
975 CGAGGTCGCAAA...TCGGGCTCCAAAGTCACCGGAACCTGCTGGCCA 1021

35 rGtyrSerAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspGly 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1022 AGTACACGGAGCTGCTCTCCGCAAGCAGCAGCAGCGAGTGAGGAGCG 1071

52 AspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPheG 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1072 GAGCTCGAAGACGCTGACGAGATT.....ATGACTGTGTTCAA 1112

68 uGlnValProAsnAsnLysGluPheLys 77
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1113 GTACATCCAGGACGAAGACGTGTTCAA 1140

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25440
seq_documentation_block:
; Sequence 25440, Application US/603600039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25440
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-60-360-039-25440
```

```
alignment_scores:
Quality: 71.50 Length: 63
Ratio: 1.663 Gaps: 4
Percent Similarity: 68.254 Percent Identity: 36.508

alignment_block:
US-09-528-682-3 x US-60-360-039-25440 ..
Align seg 1/1 to: US-60-360-039-25440 from: 1 to: 2304

53 LeuPheGlyThrValAsnGlnSerAsnPheProMetSerThrPheGlu 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1312 TTGTTCCCAACAACAACATTTGTTCTCTCCCTCAAAACCAAGGACATCAG 1361

69 nValProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLys 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1362 AATTAAACGAACAACAAATTCAACAAGGCGCACCTCTCTGCTGAA...GAAT 1408

86 yrAsp.....MetAsnPheLysLysLeuLeuArgPheMetGlu 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1409 ATGAAAAATTCATCAATCTCAAAATGAAAGTCAATCAGATTCACAGAA 1458

99 AspAspPheIleGly.....ValHisGlyGlu 107
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1459 GAA.....ATTGGTTGGATGCTTTAGTCCACGGTGAA 1491
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-284619
seq_documentation_block:
; Sequence 284619, Application US/10027632
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284619
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-284619
```

```
alignment_scores:
Quality: 70.50 Length: 81
Ratio: 1.720 Gaps: 3
Percent Similarity: 50.617 Percent Identity: 27.160

alignment_block:
US-09-528-682-3 x US-10-027-632-284619 ..
Align seg 1/1 to: US-10-027-632-284619 from: 1 to: 581
```

```
2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 TTCTTCAGAGCCAGCATCCCAAGCCACAGCAAGCTCTGAGTATTTCCT 95

18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 TTCTTCAGTCAACCCACATAATGCTTCAGGGCCTTCTGGGGAAGGCCT. 144

32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMet 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 ..GAAGGAAGATTTACAGCATACATCTTGGCAGCATTTGAAGGC..... 186

49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 .....TTCACTCTTCC 197

65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 TCAAGGATCCCAATCTCCCTCAGTCAAGAAGCTCCAGGTATC 240
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-286485
seq_documentation_block:
; Sequence 286485, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```



```
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286485
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-286485
```

```
alignment_scores:
  Quality: 70.50      Length: 81
  Ratio: 1.720        Gaps: 3
  Percent Similarity: 50.617  Percent Identity: 27.160

alignment_block:
US-09-528-682-3 x US-10-027-632-286485 ..
```

Align seg 1/1 to: US-10-027-632-286485 from: 1 to: 589

```
2 PhePheThrArgAlaLeuGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
46 TTCTTCAGAGCCAGCATCCACAGCCACAGAGCTCTGAGTATTCCCT 95
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
: : : : : ||||| ||| : : : : :
96 TTCTTCAGTCACCCACATAAATGGCTTCAGGGCCTTCTGGGAAGGCCT. 144
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMet 48
: : : : : ||||| ||| : : : : :
145 ..GAAGGAGATTTCAGCATACACTTGTGGCAGCATTTGAAGGC..... 186
```

```
49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
||| : : :
187 .....TTCACCTCTTC 197
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||| ||| : : : : : ||| : : :
198 TCAAGGGATCCCAATCTCCCTCAGTCAGTCAAGAAGCTCCAGGTATC 240
```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-203709

```
seq_documentation_block:
; Sequence 203709, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203709
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203709
```

```
alignment_scores:
  Quality: 70.50      Length: 81
  Ratio: 1.720        Gaps: 3
  Percent Similarity: 50.617  Percent Identity: 27.160

alignment_block:
US-09-528-682-3 x US-10-027-632-203709/rev ..
```

Align seg 1/1 to reverse of: US-10-027-632-203709 from: 1 to: 597

```
2 PhePheThrArgAlaLeuGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
266 TTCTTCAGAGCCAGCATCCACAGCCACAGAGCTCTGAGTATTCCCT 217
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
: : : : : ||||| ||| : : : : :
216 TTCTTCAGTCACCCACATAAATGGCTTCAGGGCCTTCTGGGAAGGCCT. 168
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMet 48
: : : : : ||||| ||| : : : : :
167 ..GAAGGAGATTTCAGCATACACTTGTGGYAGCATTTGAAGGC..... 126
49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65
||| : : :
125 .....TTCACCTCTTC 115
65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79
||| ||| : : : : : ||| : : :
114 TCAAGGGATCCCAATCTCCCTCAGTCAGTCAAGAAGCTCCAGGTATC 72
```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-286486

```
seq_documentation_block:
; Sequence 286486, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286486
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
```



US-10-027-632-286486

alignment\_scores:  
Quality: 70.50 Length: 81  
Ratio: 1.720 Gaps: 3  
Percent Similarity: 50.617 Percent Identity: 27.160

alignment\_block:  
US-09-528-682-3 x US-10-027-632-286486 ..  
Align seg 1/1 to: US-10-027-632-286486 from: 1 to: 680

2 PhePheThrArgAlaLeuGlnGlnAlaTyrProIleGluValAsnTh 18  
||||| : : : : : ||||| : : : : :  
46 TTCTTCAGCCAGCATCCACCAAGCCAGAGCTCTGAGTATTTCCCT 95  
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32  
:  
96 TTCTTCAGTCACCCACATAAATGGCTTCAGGCCCTTCTGGGAAGGCCT. 144  
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGlyMet 48  
:  
145 ..GAAGGAGATTATACAGCATACACTGTGGCAGCATTGAAGGC..... 186  
49 GlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSe 65  
||| ||| : : : : : |||||  
187 .....TTCACTCTCC 197

65 rThrPheGluGlnValProAsnAsnLysGluPheLysGlyVal 79  
||| ||| : : : : : |||||  
198 TCAAGGATCCATCTCCCTCAGTCAGTCAAGAAGCTCCAGGTATC 240

seq\_name: /cgn2\_6/ptodata/1/pna/US60-360-039-25495

seq\_documentation\_block:  
; Sequence 25495, Application US/603600039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 25495  
; LENGTH: 4770  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-60-360-039-25495

alignment\_scores:  
Quality: 70.00 Length: 101  
Ratio: 1.250 Gaps: 4  
Percent Similarity: 55.446 Percent Identity: 25.743

alignment\_block:  
US-09-528-682-3 x US-60-360-039-25495 ..  
Align seg 1/1 to: US-60-360-039-25495 from: 1 to: 4770

21 ValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArg..... 35  
||| :  
1732 GTAGATAGTGATCATGATGATGTTTCACTCCATGTCGGTAAACGTAA 1781  
36 .....TyrSerAsnPheA 40  
||| : : : : : |||||  
1782 TGTAAGTACCAGTCAACGGGATACATTGACTCCAATGAGATCATCATCA 1831

40 laLeuIleSerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThr 56  
: : : : : ||||| :  
1832 GTACAGTCAAT...GAGAACGATATGGAATAATTTCTCAGTCTTAGGTCCA 1878  
57 ValAsnGlnSerAsn.....PheProMetSe 65  
||||| :  
1879 AGAATAAGTGTAAATTCGTGCTACACCAAGGACTTCAATACAAAATTC 1928  
65 rThrPheGluGlnVal...ProAsnAsnLysGluPheLysGlyValIleS 81  
||||| : : : : : ||||| :  
1929 TACITTTGGAAGATTTTTCACCGTCCACAAAAATTTTAAGTCAGCTAAAT 1978  
81 erAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPheMet 97  
|| :  
1979 CGATTACGAAATGTTGATGTGGAATTTCTCGAAATTTTAAAGCATGTT 2028  
98 Glu 98  
: : : : :  
2029 CAG 2031

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-540-210B-6689

seq\_documentation\_block:  
; Sequence 6689, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullany, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204  
; PRIOR FILING DATE: August 1, 1996  
; PRIOR APPLICATION NUMBER: 08/903,471



```
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 6689
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00285954
; NAME/KEY: unsure
; LOCATION: 228, 243
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-210B-6689

alignment_scores:
    Quality: 69.00      Length: 48
    Ratio: 2.156       Gaps: 2
    Percent Similarity: 66.667      Percent Identity: 35.417

alignment_block:
US-09-528-682-3 x US-09-540-210B-6689/rev ..
```

```
Align seg 1/1 to reverse of: US-09-540-210B-6689 from: 1 to: 261

2 PhePheThrArgAlaLeuGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
156 TTCTTCAGAGCCAGCATCCCAAGCCACAGAGCTCTGAGTATTTCCT 107
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
: : : : : ||||| : : : : :
106 TTCTTCAGTACCCACACATAATGGCTTCAGGGCCTTCTGGGAGGCCT. 58
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly 46
: : : : : ||||| : : : : :
57 ..GAAGGAGAGATTACAGCATACACTTGTGCAGCATTTGAAGGC 16

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-34079

seq_documentation_block:
; Sequence 34079, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34079
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-10-027-632-34079

alignment_scores:
    Quality: 69.00      Length: 48
    Ratio: 2.156       Gaps: 2
    Percent Similarity: 66.667      Percent Identity: 35.417

alignment_block:
US-09-528-682-3 x US-10-027-632-34079/rev ..

Align seg 1/1 to reverse of: US-10-027-632-34079 from: 1 to: 781

2 PhePheThrArgAlaLeuGlnAlaTyrGluProIleGluValAsnTh 18
||||| : : : : : ||||| ||| : : : : :
157 TTCTTCAGAGCCAGCATCCCAAGCCACAGAGCTCTGAGTATTTCCT 108
18 rAsnThrValThrGlnIleAsnGlySerAsnGluVal.....ProL 32
: : : : : ||||| : : : : :
107 TTCYTTCAGTACCCACACATAAATGGCTTCAGGGCCTTCTGGGAGGCCT. 59
32 euAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly 46
: : : : : ||||| : : : : :
58 ..GAAGGAGAGATTACAGCATACACTTGTGCAGCATTTGAAGGC 17

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-34080
```



```

seq_documentation_block:
; Sequence 34080, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34080
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-34080

alignment_scores:
Quality: 69.00 Length: 48
Ratio: 2.156 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 35.417

alignment_block:
US-09-528-682-3 x US-10-027-632-34080/rev ..
Align seg 1/1 to reverse of: US-10-027-632-34080 from: 1 to: 781

2 PhePheThrArgAlaLeuGlnAlaTyrrGluProIleGluValAsnTh 18
||||| :|||:||||| ||| :|||:|||||
157 TTCTTCAGCGCCATCCCAACGAAGCAGACGCTCTGAGTATTTCCT 108
18 rAsnThrValThrGlnIleasnGlySerAsnGluVal.....ProL 32
107 TTCYTTCAGTCCCACATATAAATGGCTTCAGGGCCCTTCGGGAAGGCCT. 59
32 euAspGlyArgTyrrSerAsnPheAlaLeuIleSerAlaGluGly 46
58 ..GAAGAAGATTACAGCATACACTGTGGCAGCATTCGAAGC 17

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-10421-2010
seq_documentation_block:
; Sequence 2010, Application PC/TUS0210421
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Siqing
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Gaiger, Alexander
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF CANCER
; FILE REFERENCE: 210121.565PC
; CURRENT APPLICATION NUMBER: PCT/US02/10421
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2959
; SEQ ID NO 2010
; LENGTH: 394
; TYPE: DNA

```



92 LysLeuLeuArgPheMetGluAsp 100  
:::|||||  
178 TCACATG.....TTTATGAGATGAG 198

seq\_name: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:US-60-360-039-32947

seq\_documentation\_block:

; Sequence 32947, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 32947

; LENGTH: 828

; TYPE: DNA

; ORGANISM: Xylella fastidiosa

US-60-360-039-32947

alignment\_scores:

Quality: 67.50

Ratio: 2.045

Percent Similarity: 55.932

Length: 59

Gaps: 2

Percent Identity: 32.203

alignment\_block:

US-09-528-682-3 x US-60-360-039-32947 ..

Align seg 1/1 to: US-60-360-039-32947 from: 1 to: 828

2 PhePheThrArgAlaLeuGlnAlaTyrGluProIleGluValAsnTh 18

|||||  
187 TTTTTCACCGATGACTTAACCGCGGTACGCCCTGATCCGAATCC 236

|||||  
18 AsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGly 35

|||||  
237 GGACACATGCTC.....ATGCCAGCGGACGGAC 265

|||||  
35 rGtyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMetGlnAspGly 51

|||||  
266 GC.....ATCAGCCAACTCGGGCCCAATCGCGAAGGG 297

|||||  
52 AspLeuPheGlyThrValAsnGlnSer 60

|||||  
298 CCATCTTTCAAGCCCAAGGCCAGTCC 324

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-919-002-12794

seq\_documentation\_block:

; Sequence 12794, Application US/09919002

; GENERAL INFORMATION:

; APPLICANT: Leshkowitz, Dana

; APPLICANT: Liu, Jin

; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA

; FILE REFERENCE: 20411-752CON1

; CURRENT FILING DATE: 2001-07-30

; PRIOR FILING DATE: 1999-07-22

; PRIOR FILING DATE: 1998-02-13

; NUMBER OF SEQ ID NOS: 13203

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 12794

; LENGTH: 1810  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-919-002-12794

alignment\_scores:

Quality: 67.50

Ratio: 1.089

Percent Similarity: 59.048

Length: 105

Gaps: 6

Percent Identity: 27.619

alignment\_block:

US-09-528-682-3 x US-09-919-002-12794/rev ..

Align seg 1/1 to reverse of: US-09-919-002-12794 from: 1 to: 1810

11 TyrGluProIleGluValAsnThrValThrGlnIleAsnGlySe 27

|||||  
1229 TATGAAGCATCAGTCTCACCATGAATCC...AAACTCCTGAATGGTTC 1183

|||||  
27 rAsnGluValProLeuAspGly..... 34

|||||  
1182 CCAGCGGTGGTGGAGCGGCTAATCTCTGACCAGAGTGTGAGGACC 1133

|||||  
35 .....ArgTyrSerAsnPheAlaLeuIleSerAlaGluGly.....Gly 47

|||||  
1132 TGCAGAGACTGACCAATGTGGCAGCAACCTCAGGAGATGCTACCGGGGT 1083

|||||  
48 MetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMe 64

|||||  
1082 CAGACTCTCCACAT.....ACTCCCAATGAAAGTTCTATGGTGT 1042

|||||  
64 tSerThrPheGluGlnValProAsnAsnLysGluPheLysGlyValIle. 80

|||||  
1041 CACTGTCTTCAAGCCCTCAAGCTGGGCAAGCAAGCAAGTCTCTCTCC 992

|||||  
81 ..SerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPhe 96

|||||  
991 AGAGTGCCACCTGTACTACACGTGACG...GAGAAGTGCGGGCGCATC 945

|||||  
97 MetGluAspPhe 101

|||||  
944 ATGGAGTCTCTACTTC 930

seq\_name: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:US-60-360-039-25839

seq\_documentation\_block:

; Sequence 25839, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 25839

; LENGTH: 2292

; TYPE: DNA

; ORGANISM: Schizosaccharomyces pombe

US-60-360-039-25839

alignment\_scores:

Quality: 66.50

Ratio: 1.108

Percent Similarity: 57.692

Length: 104

Gaps: 5

Percent Identity: 25.962

alignment\_block:















29 uValProLeuAspGlyArgTyrSerAsnPheAlaLeuIleSerAlaGluG 46















Ratio: 1.198 Gaps: 6  
Percent Similarity: 50.962 Percent Identity: 27.885  
alignment\_block:  
US-09-528-682-3 x US-09-540-209B-1978 ..  
Align seg 1/1 to: US-09-540-209B-1978 from: 1 to: 2091  
5 ArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAsnThrVa 21  
646 AGGTTCACGACGAAGGAAGTATCCGAGGCTGAATCGGATTTAAATCACCC 695  
21 lThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgTyrSera 38  
696 GACACATTTGAATGCTAATAAATGCC.....GGAAACTATATTA 733  
38 snPheAlaLeuIle.....SerAlaGluGlyGlyMetGln 49  
734 ACCGTGCTTTGGCAGCGTTCCACCGAACAATCTGCGCGTGCTATGAGT 783  
50 AspGlyAspLeuPheGlyThrValAsnGlnSerAsnPheProMetSerTh 66  
784 GACTATGACTGGCTCTCCATATGATCGAACAACCTTT...ATCGGACA 830  
66 rPheGluGlnValProAsnAsnLysGluPheLysGlyValIleSerAla 83  
831 CTATAAC.....CGGGCTTGCTCCGGGCGC 856  
83 snValLysTyrAspMetAsnPheLysLeuLeuArgPheMetGluAsp 99  
857 AGGTAGGGGACGACAAAC.....CGTGCTATCGAGGAC 888  
100 ...AspPheile 102  
889 TTTGATTTTGTA 900  
seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-947-953-1

seq\_documentation\_block:  
; Sequence 1, Application US/09947953  
; GENERAL INFORMATION:  
; APPLICANT: DONAHUE, J. KEVIN  
; TITLE OF INVENTION: MARBAN, EDUARDO  
; FILE REFERENCE: 71699/56415  
; CURRENT APPLICATION NUMBER: US/09/947,953  
; CURRENT FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/230,311  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/295,889  
; PRIOR FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2985  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-947-953-1

alignment\_scores:  
Quality: 63.50 Length: 50  
Ratio: 2.352 Gaps: 1  
Percent Similarity: 54.000 Percent Identity: 34.000  
alignment\_block:  
US-09-528-682-3 x US-09-947-953-1 ..  
Align seg 1/1 to: US-09-947-953-1 from: 1 to: 2985

34 GlyArgTyrSerAsnPheAlaLeuIleSerAlaGluGly..... 46  
1462 GGGCGTCACAGCCATCATCTTCTGCGTAGCCTTGAGCGGCTATGACTTGG 1511

47 .....GlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlns 60  
1512 TGCTAGCTGAGGACGAGGATGTCAGATATTTACCCCGACGACCGCA 1561  
60 erAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGluPhe 76  
1562 TGCATGAGAGCATGAAGCTATTTCGATAGCATCTGCAACAACAAGTGGTTC 1611  
seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-105-299-1227  
seq\_documentation\_block:  
; Sequence 1227, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: ROSEN, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1227  
; LENGTH: 3371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-1227

alignment\_scores:  
Quality: 63.50 Length: 121  
Ratio: 1.114 Gaps: 5  
Percent Similarity: 47.107 Percent Identity: 23.967  
alignment\_block:  
US-09-528-682-3 x US-10-105-299-1227 ..  
Align seg 1/1 to: US-10-105-299-1227 from: 1 to: 3371

2 PhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnTh 18  
1824 TTCTACCCCTCGGCTCTACCTTTGACAAAGTCTCCCGTTGAG...AGTAC 1870  
18 rAsnThrValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyA 35  
1871 TACCGAACCCACGACGATTCGAGCCCTCTGAAGAG.....C 1905  
35 rGlyrSerAsnPheAlaLeuIleSerAlaGluGlyGlyMet..... 48  
1906 GTCTAAGCAATGGGATATATATTTACTGCGAGAATGGGCTCAACCTCTTC 1955  
49 .....GlnAspGlyAspLeuPheGlyThrValas 58  
1956 CTCTGGGTGGGACGACGCTCCACACAGGCT.....GTGT 1990  
58 nGlnSerAsnPheProMetSerThrPheGluGlnVal..... 70  
1991 CCAGAGCCCTTTTCAGCGCTCTCCCTTCAGTCAGATCACCAGTGGTTGA 2040  
71 .....ProAsnAsnLysGluPheLysGlyVal 79  
2041 GTGTTCTGCCAGTTCTGGATAATCCACTGTGCCAAGAAGTTTCGAGGCGTC 2090  
80 IleSerAlaAsnValLysTyrAspMetAsnPheLysLysLeuLeuArgPh 96  
2091 ATTGATAGCTTACGGGCACAGAGATCCCGGTACATGAAGCTTACCGTGGT 2140  
96 eMetGluAspAsp 100  
2141 GAAACAGGAGAC 2153

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-634-754C-766



```
; seq_documentation_block:
; Sequence 766, Application US/09634754C
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au, Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DETECTING EXPRESSION OF GENES ENCODING SECRETED PROTEINS
; FILE REFERENCE: PA-0017 US
; CURRENT APPLICATION NUMBER: US/09/634,754C
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 809
; SOFTWARE: PERL Program
; SEQ ID NO 766
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 2121260CT1
; PUBLICATION INFORMATION:
US-09-634-754C-766

alignment_scores:
  Quality: 63.00      Length: 35
  Ratio: 2.739       Gaps: 0
  Percent Similarity: 65.714      Percent Identity: 40.000

alignment_block:
US-09-528-682-3 x US-09-634-754C-766/rev ..
Align seg 1/1 to reverse of: US-09-634-754C-766 from: 1 to: 524

45 GluGlyMetGlnAspGlyAspLeuPheGlyThrValAsnGlnSerAs 61
|||||:||||| ||| ||| |||:|||||:|||||
185 GAAGGTTCCCTTCGATGATCCATTAAATGGAAGAGTAACCCAGAAAA 136

61 nPhePrometSerThrPheGluGlnValProAsnAsnLysGluPheLysG 78
||||| ||| |||:||||| ||| |||
135 CATCATGAATCCAAGAATCAGTGGATCCAGCACACAAAGAGCAAGG 86

78 lyVal 79
|||||
85 GAAT 81

seq_name: /cgn2_5/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-168431

seq_documentation_block:
; Sequence 168431, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168431
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-168431

alignment_scores:
  Quality: 63.00      Length: 26
  Ratio: 3.500       Gaps: 0
  Percent Similarity: 69.231      Percent Identity: 46.154

alignment_block:
US-09-528-682-3 x US-10-027-632-168431/rev ..
Align seg 1/1 to reverse of: US-10-027-632-168431 from: 1 to: 832

26 GlySerAsnGluValProLeuAspGlyArgTyrSerAsnPheAlaLeuI 42
|||||:||||| |||:|||||:||||| |||
269 GGCTCAAAACCGCTGCTCTATTGATGGTGCCATGCTGCCTTCCAACATGT 220

42 eSerAlaGluGlyGlyMetGlnAspGly 51
||||| |||:||||| |||
219 AAAGGATGAAGGACACACAGGATGGA 192

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-121

seq_documentation_block:
; Sequence 121, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 121
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: Misc.feature
; LOCATION: 1..1588
; OTHER INFORMATION: Ceres Seq. ID no. 1028963
US-09-935-625-121

alignment_scores:
  Quality: 63.00      Length: 47
  Ratio: 2.032       Gaps: 0
  Percent Similarity: 65.957      Percent Identity: 38.298

alignment_block:
US-09-528-682-3 x US-09-935-625-121/rev ..
Align seg 1/1 to reverse of: US-09-935-625-121 from: 1 to: 1588

56 ThrValAsnGlnSerAsnPhePrometSerThrPheGluGlnValProAs 72
|||||:|||||:||||| ||| |||:|||||:||||| |||
1110 ACAAGAGCTCAAAGCCGCTCAAGAGTCTCCTCTTCGACAAGTAATATA 1061

72 nAsnLysGluPheLysGlyValIleSerAlaAsnValLysTyrAspMetA 89
||||| |||:||||| ||| |||
1060 TGAGAGTGAAGTAACACTCATCATCAGGCTTACATCCTCTGTTCCACATA 1011

89 snPheLysLysLeuLeuArgPheMetGluAspPhePhele 102
|||||:|||||:||||| ||| |||:|||||:||||| |||
1010 ACCTTAAACAACCTTCTCGCTTCTCGAAATCATCTTCATT 970

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-330
```



```
seq_documentation_block:
; Sequence 330, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 330
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: B.fragilis
US-09-540-209B-330

alignment_scores:
  Quality: 63.00      Length: 122
  Ratio: 1.000       Gaps: 5
Percent Similarity: 51.639 Percent Identity: 24.590

alignment_block:
US-09-528-682-3 x US-09-540-209B-330 ..
Align seg 1/1 to: US-09-540-209B-330 from: 1 to: 1809

      2 PhepethrargAlaLeuGlnGlnAlaTyrcGluProIleGluValAsnTh 18
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1524 TACATTTCGGCGTAGGCTAATCCATACATCAACCTTTCGATGTAACGG 1373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

18 rAsnThrValGlnIleAsnGlySerAsnGluValProLeuAspGlyA 35
   ||| ||| : : : : : : : : : : : : : : : : : : : : : :
1374 AAATTAGTATATGATTTCGATGTTTCAGACAAATTCGATCGGATTAG 1423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

35 rGlyTy.....SerAsnPheAlaLeuIle 42
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1424 GGTTTAATATTTTGAAGAGCGTAAATAATCTCGAATGAGGAACGATT 1473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

43 SerAlaGluGlyGlyMetGlnAspGlyAspLeu..... 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1474 AATGCACCTTCGTCTATTGTTGTCGAGAGTTACGTTTTTAATGATAAACT 1523

54 ....PheGlyThrValAsnGlnSerAsnPheProMetSerThrPheGluG 69
   ||| ||| : : : : : : : : : : : : : : : : : : : : : :
1524 GAAGTTTACAACTCAACTGTTTGCATTTGGATAAAGCATCGAAAGAAC 1573

69 lnValProAsnAsnLysGluPheLys...GlyValIleSerAlaAsnVal 84
   ||| : : : : : : : : : : : : : : : : : : : : : :
1574 AGATTGCCGATAAGAGAGATTTTCAATGCGTATATAATTTCGCAAAACAGT 1623

85 LysTy.....AspMetAsnPheLysLysLeuLeuArgPheMetGluAspAs 100
   ||||| ||| : : : : : : : : : : : : : : : : : : : : : :
1624 AAATATTGGGATCTCGCTCCCAAGCAATAAATACTTTATTTCGGGACGG 1673

100 pPheileGlyValHis 105
   ||||| |||
1674 A.....GGAGTGCAT 1683

seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:us-09-053-375B-832

seq_documentation_block:
; Sequence 832, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 832
```

```
; LENGTH: 8877
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-053-375B-832

alignment_scores:
  Quality: 63.00      Length: 89
  Ratio: 1.167       Gaps: 5
Percent Similarity: 60.674 Percent Identity: 24.719

alignment_block:
US-09-528-682-3 x US-09-053-375B-832 ..
Align seg 1/1 to: US-09-053-375B-832 from: 1 to: 8877

      16 ValAsnThrAsnThrValThrGlnIleAsnGlySerAsnGluValProLe 32
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6487 GTGAGGCTCAGGAGGTGACCATGCTGAACGGAACGCTTCACCAACCTGT 6536
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

32 uAspGlyArgTySerAsn.....PheAlaLeuIleSerA 44
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6537 GACCGGCAAGAGCTTCAGCTCGGCGAGATTATTTTAAGCTCTTCAGCG 6586
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

44 lagluGlyGlyMetGln...AspGlyAspLeuPhe..... 54
   ||:||||| : : : : : : : : : : : : : : : : : : : : : :
6587 CCTCTGGAGACATGAGGACCAATGGGACAACTACCTGTATGAAATCCAG 6636

55 ...GlyThrValAsnGlnSerAsnPheProMetSerThrPheGluGlnVa 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6637 CTCCTCTCCATCACCAGCTCTTCTTACCTGCATGCGCTGCTGCCAACAT 6686

70 1.....ProAsnAsnLysGluPheLysGlyValIleSerAlaA 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6687 CTGCCAGGTGAAGCCCAATGACCAGCATTTCCAGCAGGAAGTAGGCACCT 6736
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

83 sn.....VallysTy 86
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6737 CTGACATGACCAAGTAC 6753

seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:us-09-539-800C-13583

seq_documentation_block:
; Sequence 13583, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
```







```
1226 ..AGCAATGAGTGACATTTGGATTTGCGGACCAAGGAATTT ..... 1188
43 SerAlaGluGlyGlyMetGlnAspGlyAspLeuPheGlyThrValAsnG1 59
1187 .....CG 1186
59 nSerAsnPheProMetSerThrPheGluGlnValProAsnAsnLysGluP 76
1185 CTCAGTTTTCATGGAACAATGATGATAAGATCGAGCAAGATGTT 1136
76 heLysGlyValIleSerAlaAsn.....ValLysTyr 86
1135 TGAAGGATTGACACGAGAGAACAACTAGGCGATTCGATCGGTAAAGTAC 1086
87 AspMetAsnPheLysLysLeuLeuArgPheMetGluAspAsp 100
1085 .....AACTTTTGCTTTACAACGAGGACAAAT 1059
seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-33321
seq_documentation_block:
; Sequence 33321, Application US/603600039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33321
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Desulfotobacterium hafniese
US-60-360-039-33321
```

```
alignment_scores:
  Quality: 62.50      Length: 117
  Ratio: 1.179       Gaps: 5
  Percent Similarity: 45.299   Percent Identity: 24.786
alignment_block:
US-09-528-682-3 x US-60-360-039-33321 ..
Align seg 1/1 to: US-60-360-039-33321 from: 1 to: 1581
4 ThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValAsnThrAsnTh 20
748 AGCAATGCTGCTAGCGGCATTTTCGAAGTCTTGGATGAGTACCAAGCT 797
20 rValThrGlnIleAsnGlySerAsnGluValProLeuAspGlyArgTyr 37
798 GGAATCTGAGTTGATGGG.....GTGGAATTGGCGGA..... 831
37 erAsnPheAlaLeuIleSerAlaGluGlyMetGlnAspGlyAspLeu 53
832 .....GGGCGGAGGAGCGCGGTCAATTG 855
54 PheGlyThrValAsnGlnSerAsnPheProMetSerThrPheGluGlnVa 70
856 TTCAGG.....GGAGAGCAAGCGCGCTTGGCTTACATTTAAACAGGT 899
70 lProAsnAsnLysGluPheLysGlyValIleSerAlaAsnValLysTyr 87
900 GGACCTATGCTATGAGCAAGAGGA.....GAAAGAGTTTAAAGG 940
87 spMetAsnPhe..... 90
```

```
941 ATATTAACTTTTACCCTCCGCTTGGTGAACGGGTGGCTTTTGATTGGACCC 990
91 .....LysLysLeuLeuArgPheMetG1 98
991 AGTGGCGGGGAAAAAGCTCCCTTCTGCGAGATTCTGCTTAGGTTGTAGA 1040
98 u 98
1041 A 1041
seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-36604
seq_documentation_block:
; Sequence 36604, Application US/603600039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36604
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-360-039-36604
```

```
alignment_scores:
  Quality: 62.50      Length: 109
  Ratio: 1.042       Gaps: 7
  Percent Similarity: 55.046   Percent Identity: 28.440
alignment_block:
US-09-528-682-3 x US-60-360-039-36604 ..
Align seg 1/1 to: US-60-360-039-36604 from: 1 to: 1890
1 AspPhePheThrArgAlaLeuGlnGlnAlaTyrGluProIleGluValas 17
556 GACTTTTTCAGAAAAGTGAATATCGGCTCGACCTGAATCGGTGG 605
17 nThrAsnThrValThrGlnIleAsnGlySerAsn..... 28
606 G.....CACGGTAATAACTAACCTCTACTAGTGG 634
29 .....GluValProLeuAspGlyArgTyrSerAsnPheAlaLeu 41
635 TTGTGGATGCGGTTCTCTCGCTCGACCAAGACGATCTGAATGAGAAGTTG 684
42 IleSerAlaGlu.....GlyGlyMetGlnAspGlyAspLeuPh 54
685 ATTGGGTTAAAAAGGTCACCTGGCGGTGCTCGAGGAC...TCTCTCTT 731
54 e.....GlyThrValAsnGlnSerAsnPheProMetSerThrPheGluG 69
732 CGTCAATGGAGTGGCATTCAAAAAGACATTCTCCTACGCCGCTTTTGAAC 781
69 lnValProAsnAsnLysGluPheLysGly.....ValIleSerAlaAsn 83
782 AACAGCCCC.....AAGTCTTTTAAAGACCCGCAAGATCGTGTGTTGAAT 825
84 ValLysTyrAspMetAsnPheLysLys 92
826 GTGGAGTTGGAACTGAAGAGTGAAG 852
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-045-815-5
```



[illegible]



OM of: US-09-528-682-4 to: GenEmbl:\* out\_format : pfs  
Date: Jun 18, 2002 7:35 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODE=framet.p2n\_model -DEV=xlp  
-Q=cgnt2\_1/USPTO\_spool/US09528682/runat\_18062002\_082443\_7808/app\_query.fasta\_1.689  
-DB=GenEmbl -QFMR=fastap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.500  
-XGAPOP=4.000 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40 cdi -LIST=1000 -DOALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682\_@CGN1\_1\_8225 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
Query: US-09-528-682-4  
Query length: 110  
Database: GenEmbl:\*  
Database sequences: 1797656  
Database length: 187333701  
Search time (sec): 3690.420000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_ba:ECOE0XHL	+ 181.00	369.18	3.4e-12	1262	M28523 Escherichia coli (strain
gb_ba:AE000850	- 83.50	142.08	14.99	11069	AE000850 Methanobacterium the
gb_in:PFSC04095	- 81.50	146.20	8.84	4300	AL010267 Plasmodium falciparum
gb_pat:AX067456	+ 81.00	120.95	225.27	65792	AX067456 Sequence 31 from pat
gb_in:PFU05963	+ 80.50	148.95	9.12	3474	U85963 Plasmodium falciparum a
gb_in:PFMAL4P2	+ 80.50	105.63	1.1e+03	234112	AL035475 Plasmodium falcipar
gb_in:CELM4	+ 80.00	132.72	4.76	13705	AF100670 Caenorhabditis eleg
gb_hgt:AC006890	+ 80.00	105.42	1.7e+03	238195	AC006890 Caenorhabditis eleg
gb_hgt:AC107108	- 79.00	120.80	229.64	41363	AC107108 Rattus norvegicus cl
gb_pat:AX067440	- 78.00	121.93	198.63	28626	AX067440 Sequence 15 from pat
gb_hgt:CN507EFS	- 77.00	104.82	1.8e+03	135031	AL513402 Oryza sativa chromo
gb_in:TAY15795	- 76.50	148.86	6.29	957	V15795 theileria annulata spm2
gb_ba:AF160726	+ 76.50	134.25	40.93	4374	AF160726 Peanut witches'-broom
gb_hgt:AC097767	+ 76.50	110.47	864.05	72729	AC097767 Rattus norvegicus cl
gb_in:AF410153	- 76.50	104.26	1.9e+03	146454	AF410153 Swinepox virus isol
gb_ro:MM052925	- 76.00	130.69	64.56	6585	U52925 Mus musculus coagulatio
gb_ba:AF279106	+ 76.00	105.73	1.5e+03	105184	AF279106 Uncultured proteoba
gb_hgt:PFMAL13P2_0	+ 76.00	106.13	1.6e+03	110000	AL049185 Plasmodium falcipa
gb_pat:150034	+ 74.50	144.96	10.36	918	I90034 Sequence 3 from patent
gb_pat:190033	+ 74.50	142.42	14.34	1222	I90033 Sequence 1 from patent
gb_pat:190036	+ 74.50	135.25	35.98	2745	I90036 Sequence 6 from patent
gb_ba:PP2276891	+ 74.50	134.37	40.29	3032	AJ276891 Pedicoccus pentosace
gb_pl:SCYKR021W	+ 74.50	133.52	44.94	3338	Z28246 S.cerevisiae chromoso
gb_pat:190035	+ 74.50	133.10	47.43	3500	I90035 Sequence 4 from patent
gb_in:LL14324	- 74.50	111.90	719.45	38269	LL14324 Caenorhabditis elegans
gb_in:AL137058	- 74.50	99.03	3.7e+03	163390	AL137058 Human DNA sequence
gb_ba:AP003581	+ 74.50	92.33	8.9e+03	348050	AF003581 Nostoc sp. PCC 7120
gb_ba:AE008090	+ 74.00	122.32	189.05	10473	AE008090 Agrobacterium tumefa
gb_ba:AE009125	+ 74.00	122.20	191.78	10606	AE009125 Agrobacterium tumefa
gb_hgt:AC094854	- 74.00	104.77	1.8e+03	75818	AC094854 Rattus norvegicus cl
gb_pl:AP002867	+ 74.00	99.06	3.7e+03	144322	AF002867 Oryza sativa genom
gb_pl:AP002747	+ 74.00	98.85	3.8e+03	147857	AF002747 Oryza sativa genom
gb_hgt:AC027163	- 74.00	98.33	4.1e+03	156690	AC027163 Homo sapiens chromo
gb_ba:AP003600	+ 74.00	91.42	9.9e+03	341950	AF003600 Nostoc sp. PCC 7120
gb_pl:AP0148XYL	+ 73.50	137.78	26.02	1623	Z68891 A.pisi endo-1,4-beta
gb_hgt:AC103259	+ 73.50	100.21	3.2e+03	12484	AC103259 Rattus norvegicus c
gb_vl:AF325155	- 73.50	98.31	4.1e+03	193342	AF325155 Spodoptera litura n
gb_hgt:AC095438	- 73.50	95.67	5.8e+03	187586	AC095438 Rattus norvegicus c
gb_pl:AL353732	+ 73.00	95.31	6.0e+03	173211	AL353732 Human DNA sequence
gb_ba:MMU75887	+ 72.50	121.56	208.41	7955	U75887 Methanococcus maripalud

-	gb_hgt:AC095031	72.50	95.63	5.8e+03	148238	AC095031 Homo sapiens chr
-	gb_hgt:AC099184	72.50	94.58	6.6e+03	166852	AC099184 Rattus norvegicu
-	gb_hgt:AC099452	72.50	94.30	6.9e+03	172199	AC099452 Rattus norvegicu
+	gb_hgt:AC094036	72.50	93.78	7.3e+03	182583	AC094036 Rattus norvegicu
+	gb_hgt:AC104308	72.50	93.34	7.8e+03	191971	AC104308 Homo sapiens chr
+	gb_pl:TAVDAC1	72.00	137.56	26.74	1159	TT7733 T.aestivum VDAC 1 mr
-	gb_ba:AE006005	72.00	118.18	321.50	10328	AE006005 Caulobacter cresc
-	gb_ba:U39725	72.00	116.24	411.93	12844	U39725 Mycoplasma genitali
+	gb_pl:SPAC43288	72.00	112.37	677.41	19894	AC103288 Rattus norvegicus
+	gb_in:CEC06H5	72.00	105.89	1.6e+03	41299	Z92775 Caenorhabditis eleg
+	gb_hgt:AC106600	72.00	93.91	7.2e+03	159630	AC106600 Rattus norvegicu
+	gb_hgt:CEY59A8	72.00	88.52	1.4e+04	230324	Z98870 Caenorhabditis ele
+	gb_ba:AF226440	71.50	134.56	39.34	1443	AF226440 Neisseria mening
+	gb_pat:AR003712	71.50	119.95	255.93	7493	AR003712 Sequence 5 from pa
+	gb_pat:AR029067	71.50	119.95	255.93	7493	AR029067 Sequence 5 from pa
+	gb_pat:ARI26886	71.50	119.95	255.93	7493	ARI26886 Sequence 7 from pa
+	gb_pat:IG3427	71.50	119.95	255.93	7493	IG3427 Sequence 7 from pa
+	gb_ro:MU5CCVII1	71.50	119.95	255.93	7493	LO5573 Mus domesticus coagu
+	gb_hgt:AC094529	71.50	97.40	4.6e+03	95426	AC094529 Rattus norvegicu
+	gb_pr:AC092326	71.50	94.41	6.8e+03	133709	AC092326 Homo sapiens chr
+	gb_pl:AC097926	71.50	94.26	6.9e+03	135932	AC097926 Oryza sativa chr
+	gb_hgt:AC011432	71.50	93.33	7.8e+03	131044	AC011432 Homo sapiens chr
+	gb_hgt:AC010373	71.50	93.18	7.9e+03	153693	AC010373 Homo sapiens chr
+	gb_hgt:AC093526	71.50	91.53	9.8e+03	185074	AC093526 Homo sapiens chr
+	gb_hgt:AC087558	71.50	90.04	1.2e+04	218939	AC087558 Mus musculus clo
-	gb_ba:AP003009	71.50	86.15	2.0e+04	339681	AP003009 Mesorhizobium lo
+	em_hgt:inv.AC032918	71.00	136.38	31.15	1042	AC032918 Giardia intestinal
+	gb_pat:AR083442	71.00	126.16	115.51	3300	AR083442 Sequence 6 from pa
+	gb_pat:ARI44911	71.00	126.16	115.51	3300	ARI44911 Sequence 6 from pa
+	gb_pat:AX006403	71.00	126.16	115.51	3300	AX006403 Sequence 4 from pa
+	gb_pat:118647	71.00	126.16	115.51	3300	118647 Sequence 6 from pate
+	gb_vl:AF291685	71.00	124.73	138.77	3878	AF291685 Trichoplusia ni cy
+	gb_ba:AF193063	71.00	123.90	134.21	4255	AF193063 Caulobacter cresc
+	gb_ba:AE005779	71.00	118.89	293.42	7493	AE005779 Caulobacter cresc
+	gb_hgt:AC109071	71.00	113.50	585.52	13759	AC109071 Rattus norvegicu
+	gb_ba:AF062345	71.00	110.24	889.86	19883	AF062345 Caulobacter cresc
+	gb_pr:HS04488	71.00	92.25	8.9e+03	151236	AF023883 Human DNA sequen
+	gb_hgt:AC097684	71.00	91.22	1.0e+04	169907	AC097684 Rattus norvegicu
+	gb_pr:AC006039	71.00	90.90	1.1e+04	176257	AC006039 Homo sapiens BAC
-	gb_pat:AC0094782	71.00	90.70	1.1e+04	180106	AC094782 Rattus norvegicu
-	gb_ba:AF079317	71.00	90.49	1.1e+04	184457	AF079317 Spingomonas bac
-	gb_hgt:AC091335	71.00	89.90	1.2e+04	197236	AC091335 Rattus norvegicu
-	gb_hgt:AC091335	71.00	88.29	1.5e+04	236389	AC091335 Rattus norvegicu
+	gb_in:PFMAL3P7	71.00	87.68	1.6e+04	253303	AL034559 Plasmodium falci
+	gb_pl:CCUX1ANAS	70.50	131.29	59.81	1640	LI3596 Cochliobus carbonum
+	gb_pat:HTU238895	70.50	129.57	74.56	1991	AJ238895 Helminthosporium t
+	gb_pat:128324	70.50	125.15	131.52	3280	AI01449 Recombinant DNA comp
+	gb_ba:STAFNBP	70.50	125.15	131.52	3280	I28324 Sequence 4 from pate
+	gb_ba:WSU131242	70.50	124.98	134.35	3342	J04151 S.aureus fibronectin
+	gb_ba:RT11467	70.50	121.04	222.80	5215	AJ131242 Wolinella succinog
+	gb_ba:ID0917	70.50	120.04	253.20	5836	LI1467 T.thermophilus yqjy
+	gb_ba:ID0917	70.50	108.63	1.1e+03	21131	AE008771 Salmonella typhim
+	em_hgt:inv.AC032900	70.00	137.04	28.60	760	AC032900 Giardia intestinali
+	gb_hgt:AC099206	70.00	134.93	37.48	964	AL399497 T3 end of clone AS0
+	gb_hgt:AC099206	70.00	134.06	41.93	1084	AC092546 Giardia intestinal
+	gb_ba:AF237557	70.00	128.98	80.42	1887	AF004832 Vibrio mimicus ATC
+	gb_ba:AF237557	70.00	126.13	115.98	2604	AF237557 Porphyromonas ging
+	gb_pat:AX250677	70.00	124.26	147.29	3213	AX250677 Sequence 4 from pa
+	gb_pl:AC073919	70.00	94.57	6.6e+03	91589	AC007188 Arabidopsis thali
+	gb_in:AC084459	70.00	94.45	6.7e+03	92783	AC084459 Caenorhabditis br
-	gb_hgt:PFMAL4P1_0	70.00	92.94	8.2e+03	110000	AL034557 Plasmodium falci
-	gb_pl:CPU30821	70.00	91.09	1.0e+04	135599	U03821 Cyanophora paradoxa
+	gb_pr:AL512644	70.00	90.56	1.1e+04	143919	AL512644 Human DNA sequen
+	gb_hgt:AC073919	70.00	90.00	1.2e+04	153282	AC073919 Mus musculus chr
+	gb_hgt:AC087100	70.00	88.98	1.4e+04	172116	AC087100 Mus musculus clo
+	gb_pl:AF083031	70.00	88.87	1.4e+04	174133	AF083031 Guillardia theta
+	gb_hgt:AC005505	70.00	87.96	1.5e+04	192929	AC005505 Plasmodium falci
+	gb_hgt:AC106130	70.00	87.67	1.6e+04	193342	AC106130 Rattus norvegicu
-	gb_hgt:AC105137	70.00	87.18	1.7e+04	210651	AC105137 Homo sapiens chr



gb_htg:AC005139	70.00	85.45	2.1e+04	256172	AC005139 Plasmodium falcipar	gb_in:AC023708	69.00	86.68	1.8e+04	175353	AC023708 Drosophila melan
gb_pl:AP003583	70.00	82.95	2.9e+04	339650	AP003583 Nostoc sp. PCC 7120	gb_in:AC009725	69.00	86.62	1.8e+04	176509	AC009725 Drosophila melan
gb_ba:HVLP60	69.50	135.93	32.99	764	X15691 Barley mRNA for chloro	gb_in:AC010211	69.00	86.55	1.9e+04	177857	AC010211 Drosophila melan
gb_pl:AC086020	69.50	134.44	39.94	904	AX086020 Sequence 17 from Paten	gb_in:AC007825	69.00	86.49	1.9e+04	179171	AC007825 Drosophila melan
gb_pr:AK057012	69.50	132.64	50.28	1107	AK057012 Homo sapiens cDNA FLJ	gb_in:AC009357	69.00	86.48	1.9e+04	179375	AC009357 Drosophila melan
gb_ba:AF005686	69.50	131.52	58.10	1257	AF005686 Shewanella alga DNA g	gb_in:AC008224	69.00	86.37	1.9e+04	181636	AC008224 Drosophila melan
gb_pl:AF0252086	69.50	129.68	73.51	1546	AF0252086 Mucor circoscedus epide	gb_in:AC008224	69.00	86.37	1.9e+04	181636	AC008224 Drosophila melan
gb_ba:AF290090	69.50	124.20	148.46	2869	AF290090 Staphylococcus epider	gb_in:AC008211	69.00	86.35	1.9e+04	181905	AC008211 Drosophila melan
gb_pl:AF258809	69.50	119.96	255.87	4631	AF258809 Lycopersicon esculent	gb_in:AC010051	69.00	86.23	1.9e+04	184416	AC010051 Drosophila melan
gb_ba:AF054600	69.50	115.27	466.42	7853	AF054600 Lactococcus lactis mc	gb_in:AC009090	69.00	86.14	2.0e+04	186285	AC009090 Drosophila melan
gb_pl:AF258813	69.50	113.33	598.23	9775	AF258813 Lycopersicon esculent	gb_pr:AL353734	69.00	86.13	2.0e+04	186608	AL353734 Human DNA sequen
gb_ba:AE004032	69.50	107.46	1.3e+03	18959	AE004032 Xylella fastidiosa 9	gb_in:AC007418	69.00	86.12	2.0e+04	186690	AC007418 Drosophila melan
gb_pl:AC004082	69.50	106.57	1.4e+03	20956	AC004082 Xylella fastidiosa 9	gb_in:AC014142	69.00	86.12	2.0e+04	186753	AC014142 Drosophila melan
gb_htg:YSC8H167	69.50	101.10	2.9e+03	38868	U14913 Saccharomyces cerevisi	gb_in:AC008331	69.00	86.11	2.0e+04	186938	AC008331 Drosophila melan
gb_htg:AC095872	69.50	87.95	1.6e+04	171378	AC095872 Rattus norvegicus c	gb_in:AC009848	69.00	85.98	2.0e+04	186964	AC009848 Drosophila melan
gb_htg:AC0979529	69.50	86.18	1.9e+04	209301	AC0979529 Mus musculus clone	gb_in:AC010032	69.00	85.86	2.0e+04	192393	AC010032 Drosophila melan
gb_ba:CNSPAX06	69.50	84.08	2.5e+04	265118	AJ248288 Pyrococcus abyssi c	gb_in:AE003811	69.00	83.21	2.8e+04	259445	AE003811 Drosophila melan
gb_pat:AX041922	69.50	84.08	2.5e+04	265118	AX041922 Sequence 817 from F	gb_ba:CJ1118X4	69.00	82.46	3.1e+04	282183	AL139077 Campylobacter je
gb_ba:AP000996	69.50	83.64	2.7e+04	278604	AP000996 Thermoplasma volcan	gb_in:AC003601	69.00	81.81	3.4e+04	303626	AE003601 Drosophila melan
gb_ro:AC004335	69.00	135.12	36.58	742	D10393 Rattus norvegicus mrna	gb_in:AE003477	69.00	81.78	3.4e+04	304633	AE003477 Drosophila melan
gb_pr:BC0041217	69.00	131.30	59.73	1142	BC004335 Homo sapiens, hypothe	gb_in:AE003437	69.00	81.77	3.4e+04	305018	AE003437 Drosophila melan
gb_pr:BC003605	69.00	129.82	72.24	1350	AC001217 Homo sapiens cDNA FLJ	gb_in:AE003843	69.00	81.54	3.5e+04	313115	AE003593 Drosophila melan
gb_pr:BC003430	69.00	129.67	73.58	1372	BC003605 Homo sapiens, Ssimilar	gb_in:AE003843	69.00	81.44	3.6e+04	316640	AE003843 Drosophila melan
gb_in:AY061584	69.00	127.57	96.40	1740	BC003430 Mus musculus, Ssimilar	gb_pl:CCU12314	68.50	128.85	81.78	1335	U12314 Cenchrus ciliaris cl
gb_pl:AY061584	69.00	127.28	100.06	1798	AY061584 Drosophila melanogast	gb_pat:AX023760	68.50	122.04	195.77	2877	AX023760 Sequence 17 from P
gb_pl:AY062525	69.00	126.02	117.51	2071	AY062525 Arabidopsis thaliana	gb_ro:AF284003	68.50	122.00	196.77	2890	AF284003 Glaucocystis volans B
gb_pat:AY009418	69.00	125.18	130.95	2278	AR009418 Sequence 1 from patent	gb_vl:HHGMIEXE	68.50	121.74	203.44	2976	L21763 Human herpesvirus 6
gb_ro:115623	69.00	125.18	130.95	2278	116523 Sequence 1 from patent	gb_vl:HHGMIEXF	68.50	121.74	203.44	2976	L21764 Human herpesvirus 6
gb_ro:MUSTRG1A	69.00	125.18	130.95	2278	116523 Sequence 1 from patent	gb_ba:YP2277625	68.50	121.22	217.73	3159	AJ277625 Yersinia pestis ya
gb_ba:ECHELYG	69.00	122.75	178.87	2997	X79839 E.coli O157 H7 (EDL 933	gb_vl:HHGIMWEA	68.50	118.17	321.92	4456	AJ4574 Human herpesvirus ty
gb_ba:ECU12572	69.00	121.84	201.01	3321	U12572 Escherichia coli EHEC H	gb_pl:SCDRL195W	68.50	118.06	326.28	4509	274243 S.cerevisiae chromos
gb_in:DMT1RLTR	69.00	117.89	333.50	5184	X93507 D.melanogaster retrotra	gb_pl:SCDRL089W	68.50	115.48	454.48	6035	272874 S.cerevisiae chromos
gb_ba:ECPO0AHO	69.00	113.99	549.79	8047	X860001 E.coli plasmid-DNA for	gb_ba:AE001488	68.50	102.50	819.26	10134	AE001488 Helicobacter pylor
gb_ba:AF0057695	69.00	112.77	643.28	9239	AB000801 Mycoplasma pneumoniae	gb_in:DMC190E7	68.50	102.50	2.4e+03	26100	AL020217 Drosophila melano
gb_in:DMC80H7	69.00	108.07	1.2e+03	15696	AF057695 Haemophilus ducreyi	gb_vl:SCDNAV18	68.50	101.04	2.9e+03	30752	U92288 Human herpesvirus 6
gb_in:AC005444	69.00	98.57	4.0e+03	45861	AL031027 Drosophila melanogast	gb_vl:SCDNAV18	68.50	99.80	3.4e+03	35383	X83276 S.cerevisiae DNA fo
gb_ba:AF074613	69.00	93.89	7.2e+03	77685	AC005444 Drosophila melanogast	gb_htg:AC098237	68.50	97.30	4.7e+03	46890	AC098237 Rattus norvegicus
gb_pat:AX191727	69.00	92.39	8.8e+03	92077	AF074613 Escherichia coli O15	gb_htg:AC097945	68.50	92.04	9.2e+03	84896	AC097945 Rattus norvegicus
gb_ba:AB011549	69.00	92.33	8.9e+03	92721	AX191727 Sequence 9 from Paten	gb_htg:AC098370	68.50	91.74	9.5e+03	87825	AC098370 Rattus norvegicus
gb_pat:AX191725	69.00	92.33	8.9e+03	92721	AB011549 Escherichia coli pla	gb_htg:AC019499	68.50	90.23	1.2e+04	104171	AC019499 Drosophila melan
gb_htg:AC059366	69.00	91.87	9.4e+03	97648	AX191725 Sequence 7 from Paten	gb_htg:AC093974	68.50	89.42	1.3e+04	114067	AC093974 Rattus norvegicu
gb_htg:AC059366	69.00	91.87	9.4e+03	97648	AC059366 Rattus norvegicus cl	gb_htg:AC094479	68.50	88.46	1.5e+04	127174	AC094479 Rattus norvegicu
gb_htg:AC097987	69.00	89.16	1.2e+04	117368	AC095366 Rattus norvegicus c	gb_htg:AC097430	68.50	87.63	1.6e+04	139596	AC097430 Homo sapiens chr
gb_pl:AC018848	69.00	89.13	1.3e+04	132455	AC097987 Rattus norvegicus c	gb_htg:AF002445	68.50	86.61	1.8e+04	156608	AF002445 Homo sapiens chr
gb_htg:AC010111	69.00	88.74	1.4e+04	138938	AC018848 Arabidopsis thalian	gb_vl:AB021506	68.50	86.34	1.9e+04	161573	AB021506 Human herpesviri
gb_in:AC004759	69.00	88.50	1.4e+04	142799	AC010111 Drosophila melanoga	gb_htg:AF002416	68.50	86.21	1.9e+04	163989	AF002416 Homo sapiens chr
gb_htg:OSUN00074	69.00	88.21	1.5e+04	147491	AL606641 Oryza sativa chromo	gb_in:AC104288	68.50	86.19	1.9e+04	164252	AC104288 Drosophila melan
gb_in:AC009208	69.00	87.92	1.6e+04	152443	AC009208 Drosophila melanoga	gb_pr:AC012148	68.50	86.01	2.0e+04	167619	AC012148 Homo sapiens chr
gb_in:AC008210	69.00	87.85	1.6e+04	153648	AC008210 Drosophila melanoga	gb_pr:AC090231	68.50	85.83	2.0e+04	171076	AC090231 Homo sapiens chr
gb_pr:AC005951	69.00	87.75	1.6e+04	153648	AC005951 Homo sapiens chromo	gb_htg:AC079217	68.50	85.52	2.1e+04	17205	AC079217 Mus musculus chr
gb_htg:AC098190	69.00	87.70	1.6e+04	156305	AC098190 Rattus norvegicus c	gb_htg:AC095589	68.50	85.21	2.2e+04	183551	AC095589 Rattus norvegicu
gb_in:AC093047	69.00	87.67	1.6e+04	156806	AC093047 Rattus norvegicus c	gb_htg:AC105367	68.50	84.34	2.5e+04	202388	AC105367 Rattus norvegicu
gb_in:AC093039	69.00	87.63	1.6e+04	157561	AC093039 Drosophila melanoga	gb_ba:AJ414154	68.50	84.27	2.5e+04	204050	AJ414154 Yersinia pestis
gb_in:AC09253	69.00	87.47	1.6e+04	160282	AC09253 Drosophila melanoga	gb_htg:AC090359	68.50	83.56	2.7e+04	221041	AC090359 Homo sapiens chr
gb_htg:AC094612	69.00	87.25	1.7e+04	164427	AC094612 Rattus norvegicus c	gb_htg:AP001566	68.50	83.51	2.7e+04	222408	AP001566 Homo sapiens chr
gb_htg:AC098660	69.00	87.07	1.7e+04	176698	AC098660 Rattus norvegicus c	gb_in:AE003420	68.50	80.77	3.9e+04	302855	AE003420 Drosophila melan
gb_in:AC008135	69.00	87.04	1.7e+04	168306	AC098660 Rattus norvegicus c	gb_ov:GGU68380	68.00	128.65	83.92	1211	U68380 Gallus gallus single
gb_in:AC107487	69.00	86.91	1.8e+04	170900	AC008135 Drosophila melanoga	gb_pl:CAR012751	68.00	129.59	87.87	1261	AJ012751 Cicer arietinum mR
gb_in:AC010561	69.00	86.88	1.8e+04	171439	AC107487 Drosophila melanoga	gb_ro:AF170906	68.00	127.50	97.20	1378	AF170906 Mus musculus Lck-a
gb_in:AC010561	69.00	86.88	1.8e+04	171439	AC010561 Drosophila melanoga	gb_pat:AX188133	68.00	124.68	139.63	1895	AX188133 Sequence 3828 from
gb_pr:AC007394	69.00	86.88	1.8e+04	171439	AC007394 Homo sapiens BAC cl	gb_ba:AF195009	68.00	121.88	199.96	2599	AF195009 Helicobacter pylor
gb_in:AC010668	69.00	86.84	1.8e+04	171468	AC010561 Drosophila melanoga	gb_pr:ECOHLE2	68.00	121.73	203.81	2643	BC000269 Homo sapiens, clon
gb_htg:AC006489	69.00	86.78	1.8e+04	173298	AC010668 Drosophila melanoga	gb_pr:HUMHT31A	68.00	120.48	239.41	3045	M81823 Escherichia coli hem
gb_in:AC007467	69.00	86.78	1.8e+04	173422	AC006489 Drosophila melanoga	gb_pl:SCYNL082W	68.00	120.48	239.41	3045	M81823 Escherichia coli hem
gb_in:AC009344	69.00	86.77	1.8e+04	173634	AC007467 Drosophila melanoga	gb_pr:BC009213	68.00	119.69	264.68	3326	271358 S.cerevisiae chromos
gb_in:AC008342	69.00	86.70	1.8e+04	175007	AC009344 Drosophila melanoga	gb_pl:YSCPM51A	68.00	119.47	275.58	3336	AC09213 Homo sapiens, clon
gb_in:AC010576	69.00	86.69	1.8e+04	175192	AC008342 Drosophila melanoga	gb_pl:SCYNL081C	68.00	119.47	272.29	3410	BC29688 S.cerevisiae PMS1 ge
					AC010576 Drosophila melanoga		68.00	118.91	292.71	3634	271357 S.cerevisiae chromos



gb_pt:AF126008	4290	AF126008 Homo sapiens breast c	gb_ba:AE001580	30885	AE001580 Borrelia burgdorferi
gb_ba:WBOAPULNS	4375	L07762 Thermoanaerobacterium s	gb_ba:MLCB2548	38916	AL023093 Mycobacterium lep
gb_ba:WBOAP12073	4838	L012073 Wolbachia sp. Wri gl	gb_ba:AE001584	52971	AL001584 Borrelia burgdorferi
gb_pt:AF387101	8442	AF387101 Homo sapiens protein	gb_hgt:AC100499	58608	AC100499 Mus musculus clon
gb_ba:AB045235	9386	AB045235 Wolbachia sp. wta1 r1	gb_pt:SC2E9781	63001	U18916 Saccharomyces cerev
gb_pt:AF406992	9468	AF406992 Homo sapiens A-kinase	gb_hgt:AC106273	71188	AC106273 Rattus norvegicus
gb_pt:AB055890	10156	AB055890 Homo sapiens c-lbc n	gb_hgt:AC098320	72456	AC098320 Rattus norvegicus
gb_in:AB001407	13437	AC001407 Plasmodium falcipar	gb_hgt:AC097831	73215	AC097831 Rattus norvegicus
gb_pt:SC23CDS	31431	X86470 S. cerevisiae PWS1, TPM	gb_hgt:AC101587	73789	AC101587 Mus musculus clon
gb_pt:AF1356983	51923	AF1356983 Human DNA sequence	gb_pt:SCD9461	78500	U33007 Saccharomyces cerev
gb_hgt:AC107592	73497	AC107592 Rattus norvegicus cl	gb_ov:AL591180	80272	AL591180 zebrafish DNA seq
gb_hgt:AC090931	90280	AC090931 Homo sapiens chromos	gb_pt:AB019236	81494	AB019236 Arabidopsis thali
gb_hgt:LMFCH36_14	102000	Continuation (15 of 36) of 1	gb_hgt:AC095541	83131	AC095541 Rattus norvegicus
gb_hgt:OSJN06017	133428	AL066150 Oryza sativa chromo	gb_hgt:AC107555	84519	AC107555 Rattus norvegicus
gb_hgt:AC096100	149325	AC096100 Rattus norvegicus c	gb_ov:AL591175	89232	AL591175 zebrafish DNA seq
gb_hgt:AC094715	153223	AC094715 Rattus norvegicus c	gb_hgt:AC099305	91953	AC099305 Rattus norvegicus
gb_hgt:AC106404	157373	AC106404 Rattus norvegicus c	gb_pt:FBK7	98581	AC007727 Arabidopsis thali
gb_pt:AC1195074	160755	AX195074 Sequence 544 from F	gb_hgt:AC095010	99203	AC095010 Rattus norvegicus
gb_hgt:AC094557	160791	AC094557 Rattus norvegicus c	gb_hgt:AC084726	101375	AC084726 Mus musculus chr
gb_hgt:AC007602	160865	AC007602 Homo sapiens chromo	gb_hgt:OSJN00045	108479	AL606610 Oryza sativa chr
gb_vl:AF081810	161046	AF081810 Lymantria dispar nu	gb_hgt:AC094266	115062	AC094266 Rattus norvegicu
gb_hgt:AC080157	169801	AC080157 Rattus norvegicus c	gb_pt:AF003049	117000	AC003049 Homo sapiens gen
gb_hgt:AC105656	178982	AC105656 Rattus norvegicus c	gb_hgt:AC095745	117980	AC095745 Rattus norvegicu
gb_hgt:AC098141	180313	AC098141 Rattus norvegicus c	gb_hgt:AC103239	137897	AC103239 Rattus norvegicu
gb_hgt:AC026802	192386	AC026802 Homo sapiens chromo	gb_hgt:AC094726	143135	AC094726 Rattus norvegicu
gb_hgt:PFMAL13P6	204652	AL049183 Plasmodium falcipar	gb_hgt:AC068373	155186	AC068373 Homo sapiens chr
gb_hgt:AC096325	233382	AC096325 Rattus norvegicus c	gb_hgt:AC106473	156169	AC106473 Rattus norvegicu
gb_ba:CU11168X3	317511	AL139076 Campylobacter jejun	gb_hgt:AC094654	156476	AC094654 Rattus norvegicu
gb_pt:AX142405	639	AX142405 Sequence 1127 from Pat	gb_hgt:AC103366	156614	AC103366 Homo sapiens chr
gb_pt:EB39874	1212	E39874 Method for reducing the	gb_hgt:AC051640	158514	AC051640 Homo sapiens chr
gb_ov:DUKARGINX	1582	M35133 Anas platyrhynchos delt	gb_pt:AL499604	160796	AL499604 Human DNA sequen
gb_ba:RATGASB	1645	M18769 Rat liver beta-galactos	gb_hgt:AC103206	163486	AC103206 Rattus norvegicu
gb_ba:CACSPD	2181	Z50009 C. acetobutylicum CSPD	gb_hgt:AC094061	165502	AC094061 Rattus norvegicu
gb_in:AV075532	2395	AY075532 Drosophila melanogast	gb_hgt:AC096596	170674	AC096596 Rattus norvegicu
gb_ba:AF270077	3167	AF270077 Staphylococcus epider	gb_hgt:AC106311	171237	AC106311 Rattus norvegicu
gb_pt:AF362372	6988	AF362372 Dictyostellium discoide	gb_hgt:AC097018	173083	AC097018 Rattus norvegicu
gb_ba:AB006675	9167	AX145395 Sequence 4117 from Pa	gb_hgt:AC102971	174436	AC102971 Rattus norvegicu
gb_pt:AX353827	10732	AX056775 Sulfolobus solfataric	gb_hgt:AC087156	177630	AC087156 Mus musculus clo
gb_pt:AX353827	13288	AX353827 Sequence 27 from Pat	gb_hgt:AC099102	178544	AC099102 Rattus norvegicu
gb_in:CEY51A2A	35303	AL032635 Caenorhabditis elega	gb_pt:AC009756	178624	AC009756 Homo sapiens, cl
gb_in:AC084445	62584	AC084445 Caenorhabditis briggs	gb_hgt:AC084240	181282	AC084240 Rattus norvegicu
gb_pt:AC092113	86372	AC093113 Genomic sequence for	gb_hgt:AC106452	183818	AC106452 Rattus norvegicu
gb_pt:AFV1715	8748	AL031032 Arabidopsis thaliana	gb_hgt:AC094159	184640	AC094159 Rattus norvegicu
gb_hgt:AC017901	89670	AL017901 Drosophila melanogast	gb_hgt:AC095692	198850	AC095692 Rattus norvegicu
gb_hgt:AC020798	114979	AC020798 Mus musculus clone	gb_hgt:AC092296	203650	AC092296 Homo sapiens chr
gb_pt:AL135932	143321	AL135932 Human DNA sequence	gb_ba:AF003192	301550	AC003192 Clostridium perfr
gb_hgt:AC095893	152867	AC095893 Rattus norvegicus c	gb_ba:MLEPRTN1	344050	AL583917 Mycobacterium le
gb_hgt:AC095244	156647	AC095244 Rattus norvegicus c	em_hgt:inv:AC031479	878	AC031479 Giardia intestinali
gb_in:AC092223	171598	AC092223 Drosophila melanoga	em_hgt:inv:AC078143	900	AC078143 Giardia intestinali
gb_hgt:AC020695	182513	AC020695 Homo sapiens chromo	em_hgt:inv:AC048725	984	AC048725 Giardia intestinali
gb_hgt:AC084393	182742	AC084393 Homo sapiens chromo	em_hgt:inv:AC031137	1002	AC031137 Giardia intestinali
gb_hgt:AC098987	193893	AC098987 Rattus norvegicus c	em_hgt:inv:AC047374	1123	AC047374 Giardia intestinali
gb_hgt:AC1365178	195162	AL365178 Homo sapiens chromo	gb_pt:AF140553	1206	AF140553 Avena sativa DNA-b
gb_pt:AFCHRI570	192861	AL161584 Arabidopsis thalian	gb_pt:AB011286	1305	AB011286 Candida maltosa EP
gb_hgt:AC017086	211365	AC017086 Homo sapiens chromo	gb_ov:AF332044	2878	AF332044 Scirtus niger BRCA
gb_in:AE003641	299556	AE003641 Drosophila melanoga	gb_ov:BOVPIGR1	3527	L04797 Bos taurus polymeric
gb_in:DROSADH02	320754	AE003408 Drosophila melanoga	gb_ov:BNPIGR1	3630	X81371 B. taurus mRNA for bo
gb_sts:GL17579	400	GL17579 human STS SHGC-11666 cld	gb_pt:AF172654	3630	AF172654 Sequence 5 from pa
gb_ba:CFPOLC	448	X86524 C. perfringens folC gene	gb_pt:SCADR6	4830	X12493 Yeast AD6 gene for
gb_ba:HFU005677	1541	U05677 Helicobacter pylori 87-	gb_pt:SCU15219	5521	U15219 Saccharomyces cerevi
gb_pt:189393	1541	I89393 Sequence 3 from patent	gb_ba:ECOHLY	8211	M10133 E.coli (J96) hlyc, h
gb_pt:AF361830	1937	AF361830 Arabidopsis thaliana	gb_ba:AE006596	10029	AE006596 Streptococcus pyo
gb_pt:AV064044	2768	AY064044 Arabidopsis thaliana	gb_ba:AE009152	11049	AE009152 Agrobacterium tum
gb_pt:AV059877	2812	AY059877 Arabidopsis thaliana	gb_ba:AE009569	11225	AE009569 Complete sequence o
gb_ba:AE007843	12251	AE007843 Clostridium acetobut	gb_vl:FPV11KB	11225	D00295 Fowlpox virus genom
gb_ba:AE000598	12551	AE000598 Helicobacter pylori;	gb_ba:AE006095	11249	AE006095 Pasteurella multo
gb_in:CELC0566	15075	AF045635 Caenorhabditis elega	gb_pt:SCYBL101C	12242	Z35862 S. cerevisiae chromo
gb_ba:AE002154	19497	AE002154 Ureaplasma urealytic	gb_ba:AE008118	12606	AE008118 Agrobacterium tum
gb_in:CBRG01P23	20422	AC084476 Caenorhabditis briggs	gb_pt:U33335	35550	U33335 Saccharomyces cerev
gb_ba:AE001578	29838	AE001578 Borrelia burgdorferi	gb_vl:AL356491	49336	AL356491 faba monkey tumor
gb_ba:AE001576	30223	AE001576 Borrelia burgdorferi	gb_pt:AL356491	60350	AL356491 Human DNA sequenc
gb_ba:AE001577	30299	AE001577 Borrelia burgdorferi	gb_hgt:AC102251	65903	AC102251 Mus musculus clon
gb_ba:AE001575	30750	AE001575 Borrelia burgdorferi	gb_pt:SCDNC2	78600	X79489 S. cerevisiae genom







gb_htg:AC092921_1	-	65.50	83.35	2.8e+04	110000	Continuation (2 of 5) of AC093316 Mus musculus chromo	gb_htg:AC099125	-	65.00	79.27	4.7e+04	154600	AC099125 Rattus norvegicu
gb_htg:AC093316_0	+	65.50	83.35	2.8e+04	110000	AC093316 Mus musculus chromo	gb_htg:AC099165	+	65.00	79.11	4.8e+04	157388	AC099165 Rattus norvegicu
gb_htg:AC095684	-	65.50	83.07	3.1e+04	113543	AC095684 Rattus norvegicus	gb_htg:AC106679	+	65.00	79.08	4.8e+04	157913	AC106679 Rattus norvegicu
gb_pr:AC004930	+	65.50	82.69	3.0e+04	118451	AC004930 Homo sapiens PAC cl	gb_htg:AC084369	+	65.00	79.00	4.9e+04	159284	AC084369 Homo sapiens chr
gb_htg:HSJ08706	-	65.50	82.59	3.1e+04	119907	AL589706 Human DNA sequence	gb_htg:AC095736	+	65.00	78.92	4.9e+04	160845	AC095736 Rattus norvegicu
gb_htg:OSJN01011	+	65.50	82.51	3.1e+04	120990	AL607103 Oryza sativa chromo	gb_htg:AC106235	+	65.00	78.82	5.0e+04	160845	AC106235 Rattus norvegicu
gb_htg:AC096300	+	65.50	82.31	3.2e+04	123763	AC096300 Rattus norvegicus	gb_htg:AC103020	+	65.00	78.86	5.1e+04	165585	AC103020 Rattus norvegicu
gb_htg:AC095959	-	65.50	82.04	3.3e+04	127587	AC005959 Homo sapiens, ** s	gb_pr:AC019211	+	65.00	78.56	5.1e+04	167471	AC019211 Homo sapiens BAC
gb_htg:AC099196	+	65.50	81.57	3.5e+04	134490	AC099196 Rattus norvegicus	gb_htg:AC015549	+	65.00	78.54	5.2e+04	167849	AC105549 Rattus norvegicu
gb_pr:HSJ12208	+	65.50	81.20	3.7e+04	140211	AL096678 Human DNA sequence	gb_htg:AC104786	+	65.00	78.45	5.2e+04	169520	AC104786 Homo sapiens chr
gb_htg:AC025736	+	65.50	80.93	3.8e+04	144585	AC025736 Homo sapiens chromo	gb_htg:AC098169	+	65.00	78.41	5.2e+04	170378	AC098169 Rattus norvegicu
gb_htg:AC099260	+	65.50	80.83	3.8e+04	146099	AC099260 Rattus norvegicus	gb_htg:AC099159	-	65.00	78.30	5.3e+04	172373	AC099159 Rattus norvegicu
gb_htg:OSJN00134	+	65.50	80.82	3.8e+04	146274	AL662944 Oryza sativa chromo	gb_htg:AC021696	+	65.00	78.30	5.3e+04	172378	AC021696 Homo sapiens clo
gb_htg:PFMAL181	+	65.50	80.46	4.0e+04	152409	AL031744 Plasmodium falci-par	gb_htg:AC068626	+	65.00	78.23	5.4e+04	173742	AC068626 Homo sapiens chr
gb_htg:AC025389	+	65.50	80.17	4.2e+04	157419	AC025389 Homo sapiens chromo	gb_htg:AC098157	+	65.00	78.23	5.4e+04	175293	AC098157 Rattus norvegicu
gb_htg:AC013563	+	65.50	80.12	4.2e+04	158281	AC013563 Homo sapiens chromo	gb_htg:AC018948	+	65.00	78.15	5.4e+04	175360	AC018948 Homo sapiens chr
gb_htg:AC094457	+	65.50	79.99	4.3e+04	160791	AC094457 Rattus norvegicus	gb_pr:AC012048	+	65.00	77.97	5.5e+04	178963	AC012048 Homo sapiens chr
gb_htg:AC094073	+	65.50	79.92	4.3e+04	162000	AC094073 Rattus norvegicus	gb_htg:AC098556	+	65.00	77.95	5.6e+04	179428	AC098556 Rattus norvegicu
gb_htg:AC097029	+	65.50	79.90	4.3e+04	162408	AC097029 Rattus norvegicus	gb_htg:AC068090	+	65.00	77.47	5.9e+04	189245	AC068090 Homo sapiens chr
gb_pr:AL353621	+	65.50	79.80	4.4e+04	164115	AL353621 Human DNA sequence	gb_htg:AC106227	+	65.00	77.47	5.9e+04	189265	AC106227 Rattus norvegicu
gb_htg:AC094945	+	65.50	79.70	4.4e+04	165989	AC094945 Rattus norvegicus	gb_htg:AL671886	+	65.00	77.41	5.9e+04	190708	AL671886 Mus musculus chr
gb_htg:AC098920	-	65.50	79.16	4.8e+04	176530	AC098920 Rattus norvegicus	gb_htg:AL606922	+	65.00	77.36	6.0e+04	191782	AL606922 Mus musculus chr
gb_ro:AL590626	+	65.50	79.02	4.8e+04	179282	AL590626 Mouse DNA sequence	gb_pr:AC078851	+	65.00	76.93	6.3e+04	201303	AC078851 Homo sapiens BAC
gb_htg:AL670035	+	65.50	78.91	4.9e+04	181563	AL670035 Mus musculus chromo	gb_htg:AL645586	+	65.00	76.87	6.4e+04	202505	AL645586 Mus musculus chr
gb_htg:AL359264	+	65.50	78.91	4.9e+04	181627	AL359264 Homo sapiens chromo	gb_htg:AC094414	+	65.00	76.85	6.4e+04	203023	AC094414 Rattus norvegicu
gb_htg:AC108676	+	65.50	78.36	5.3e+04	193103	AC108676 Homo sapiens chromo	gb_htg:AL669952	+	65.00	76.84	6.4e+04	203356	AL669952 Mus musculus chr
gb_htg:AC108416	+	65.50	78.29	5.3e+04	194686	AC108416 Rattus norvegicus	gb_htg:AC092166	+	65.00	76.51	6.7e+04	203958	AC092166 Homo sapiens chr
gb_ba:BSUB00007	+	65.50	77.34	6.0e+04	216750	Z99110 Bacillus subtilis com	gb_htg:AC096324	+	65.00	75.63	7.5e+04	229956	AC096324 Rattus norvegicu
gb_htg:AL606909	+	65.50	77.04	6.2e+04	224061	AL606909 Mus musculus chromo	gb_htg:AC095680	+	65.00	75.25	7.8e+04	243167	AC095680 Rattus norvegicu
gb_pr:AC012150	+	65.50	76.92	6.3e+04	227305	AC012150 Homo sapiens 12 BAC	gb_htg:AC106538	+	65.00	74.82	8.3e+04	255336	AC106538 Rattus norvegicu
gb_ba:AP001519	+	65.50	74.35	8.8e+04	303650	AP001519 Bacillus halodurans	gb_ba:SMES91785	-	65.00	73.80	9.4e+04	286550	AL591785 Sinorhizobium me
gb_ba:AP003586	+	65.50	73.11	1.0e+05	349050	AP003586 Nostoc sp. PCC 7120	gb_ba:PMAL13P3	+	65.00	73.36	9.9e+04	301150	AP003190 Clostridium perf
gb_pr:NGA243078	+	65.00	129.48	75.46	536	AL243078 Nectria galligena part	gb_htg:PFMAL13P3	+	65.00	72.87	1.1e+05	382221	AL049184 Plasmodium falci
gb_pr:NGA243081	+	65.00	129.48	75.46	536	AL243081 Nectria galligena part	gb_htg:PFMAL13P3	+	65.00	72.79	1.1e+05	382221	AL049184 Plasmodium falci
gb_pr:AL12697	-	65.00	127.67	95.10	657	AL12697 3' terminal of Hly A gen	gb_ba:AP003592	+	65.00	72.07	1.2e+05	348250	AP003592 Nostoc sp. PCC 7
gb_ba:ERKEPCD	-	65.00	121.28	215.84	1351	M12049 Eikenella corrodens pil	gb_pr:AC093747	+	64.50	125.37	127.91	755	AC093747 Sequence 7 from pate
gb_ph:PM2AAGA	+	65.00	118.64	302.87	1820	M33144 Bacteriophage M2 primex	gb_pr:AR148418	+	64.50	125.36	127.90	756	AR148418 Sequence 7 from pate
gb_pr:AR272749	+	65.00	115.60	447.37	2565	AF272749 Zea mays kinesin heav	gb_pr:AR143089	+	64.50	124.65	140.08	819	AR143089 Sequence 1811 from
gb_in:AR271785	+	65.00	114.92	488.04	2769	AF217185 Caenorhabditis elegans	gb_pr:AR143089	+	64.50	123.11	170.79	975	AR143089 Sequence 1811 from
gb_pr:AR421953	+	65.00	110.13	901.80	4752	AF421953 Kluyveromyces lactis	gb_pr:AR143089	+	64.50	120.15	249.54	1361	AR038817 Sequence 4 from pa
gb_in:DDIMYHC	+	65.00	107.11	1.2e+03	6681	M14628 D. discoideum myosin hea	gb_pr:AR143089	+	64.50	120.15	249.54	1361	AR038817 Sequence 4 from pa
gb_ba:U39716	+	65.00	103.32	2.3e+03	10252	U39716 Mycoplasma genitalium	gb_pr:AR143089	+	64.50	119.77	262.09	1421	L34016 Borrelia burgdorferi
gb_ba:AE000889	+	65.00	102.95	2.3e+03	10688	AE000889 Methanobacterium the	gb_ba:BOBS1A	+	64.50	117.05	371.42	1931	U25157 Zea mays beta-D-gluc
gb_ba:AE001514	+	65.00	101.32	2.8e+03	12842	AE001514 Helicobacter pylori	gb_pr:ZMU25157	+	64.50	116.98	374.70	1946	U33816 Zea mays strain B73
gb_ba:AE001475	+	65.00	100.68	3.0e+03	13809	AE001475 Helicobacter pylori	gb_pr:ZMU25157	+	64.50	116.98	374.70	1946	U33816 Zea mays strain B73
gb_ba:AE001581	+	65.00	93.61	7.5e+03	30651	AE001581 Borrelia burgdorferi	gb_pr:ZMU25157	+	64.50	116.98	374.70	1946	U33816 Zea mays strain B73
gb_ba:AF389114	+	65.00	92.08	9.1e+03	36410	AF389114 Rhizobium sp. NGR234	gb_pr:ZMU25157	+	64.50	114.03	547.15	2715	X74217 Z. mays mRNA for beta
gb_in:CEC4788	+	65.00	90.33	1.1e+04	44371	Z75530 Caenorhabditis elegans	gb_ro:RNGPLV	+	64.50	113.59	578.64	2852	BD006432 Compounds and meth
gb_ph:AF47491	+	65.00	88.69	1.4e+04	53373	AF47491 Bacteriophage phiE12	gb_pr:BD006552	+	64.50	113.59	578.64	2852	BD006552 Compounds and meth
gb_htg:AC066803	+	65.00	88.58	1.4e+04	54039	AL606803 Human DNA sequence	gb_ba:AF269638	+	64.50	113.10	616.15	3014	AF269638 Staphylococcus epi
gb_htg:AC094699	+	65.00	88.54	1.4e+04	54338	AC094699 Rattus norvegicus cl	gb_pr:AX144958	+	64.50	113.10	616.15	3014	AX144958 Sequence 3680 from
gb_htg:AC107477	+	65.00	87.57	1.6e+04	60583	AC107477 Rattus norvegicus cl	gb_ba:AF269815	+	64.50	112.91	631.51	3080	AX145133 Sequence 3855 from
gb_htg:AC103472	+	65.00	86.34	1.9e+04	60789	AC103472 Rattus norvegicus cl	gb_pr:AX145133	+	64.50	112.91	631.51	3080	AX145133 Sequence 3855 from
gb_htg:AC105464	+	65.00	85.01	2.3e+04	80919	AC105464 Rattus norvegicus cl	gb_pr:YSCGSH	+	64.50	112.16	695.76	3354	M30449 Yeast (S.cerevisiae) h
gb_htg:AC105464	+	65.00	85.01	2.3e+04	80919	AC105464 Rattus norvegicus cl	gb_pr:YSCGSH	+	64.50	111.26	781.03	3713	J03139 S.cerevisiae heat sh
gb_htg:AC105464	+	65.00	84.92	2.3e+04	81710	AC105464 Rattus norvegicus cl	gb_pr:YSCGSH	+	64.50	110.77	831.46	3923	Z72596 S.cerevisiae chromos
gb_htg:AC10950	+	65.00	84.65	2.4e+04	84212	AC10950 Homo sapiens clone R	gb_pr:YSCGSH	+	64.50	110.76	832.16	3926	M22040 S.cerevisiae heat sh
gb_in:PFMAL13P5	+	65.00	83.33	2.5e+04	86827	AL034556 Plasmodium falci-par	gb_ba:HPU95371	+	64.50	110.07	908.96	4243	U95971 Helicobacter pylori
gb_htg:AC097401	+	65.00	83.33	2.8e+04	97731	AL034556 Plasmodium falci-par	gb_ba:ECU04846	+	64.50	108.84	1.1e+03	4875	U04846 Escherichia coli B87
gb_htg:AC105497	+	65.00	82.91	2.9e+04	101045	AC105497 Rattus norvegicus cl	gb_pr:HS5620	+	64.50	108.84	1.1e+03	4875	U04846 Escherichia coli B87
gb_htg:AC095879	+	65.00	82.91	3.0e+04	102529	AC095879 Rattus norvegicus	gb_vl:GM042608	+	64.50	107.71	1.2e+03	5537	U42608 Glycine max clathrin
gb_htg:AC095707	+	65.00	82.87	3.0e+04	102976	AC095707 Rattus norvegicus	gb_vl:GM042608	+	64.50	105.06	1.7e+03	7470	AF006408 Fowlpox virus stra
gb_htg:AC105535	-	65.00	82.29	3.2e+04	109924	AC105535 Rattus norvegicus	gb_pr:YSCDR1A	+	64.50	104.09	2.0e+03	8433	L19540 Saccharomyces cerevi
gb_htg:AC103580	+	65.00	82.17	3.3e+04	111442	AC103580 Rattus norvegicus	gb_in:DMU77947	+	64.50	104.02	2.0e+03	8433	L19540 Saccharomyces cerevi
gb_htg:AC106226	+	65.00	81.78	3.4e+04	116438	AC106226 Rattus norvegicus	gb_ba:AE006317	+	64.50	101.95	2.6e+03	10610	U006317 Lactococcus lacti
gb_htg:AC096206	+	65.00	81.68	3.5e+04	117774	AC096206 Rattus norvegicus	gb_ba:AE006942	+	64.50	96.10	5.3e+03	19969	AC094536 Rattus norvegicu
gb_htg:AC097548	+	65.00	81.45	3.6e+04	120838	AC097548 Rattus norvegicus	gb_ba:AE006942	+	64.50	96.10	5.3e+03	19969	AC094536 Rattus norvegicu
gb_htg:AC097912	+	65.00	81.00	3.8e+04	127080	AC097912 Rattus norvegicus	gb_pr:SPAC1486	+	64.50	92.83	8.3e+03	29676	AL133357 S.pombe chromoso
gb_pr:HSJ661G12	-	65.00	79.93	4.3e+04	134553	AL049797 Human DNA sequence	gb_pr:SC8337	+	64.50	91.49	9.8e+03	34506	Z46659 S.cerevisiae chromo
gb_pr:AC006998	+	65.00	79.88	4.3e+04	142660	AC006998 Homo sapiens BAC cl	gb_ba:MT13E10	+	64.50	91.36	1.0e+04	35019	Z95324 Mycobacterium tuber
gb_htg:AC099260	+	65.00	79.77	4.4e+04	146099	AC099260 Rattus norvegicus	gb_in:CE001084	+	64.50	91.08	1.0e+04	36170	Z50874 Caenorhabditis eleg
gb_htg:AC008691	-	65.00	79.65	4.5e+04	148138	AC008691 Homo sapiens chromo	gb_pr:AC007288	+	64.50	89.01	1.3e+04	45300	AC007288 Arabidopsis thali
gb_htg:AC094736	+	65.00	79.63	4.5e+04	148439	AC094736 Rattus norvegicus	gb_htg:AC018082	+	64.50	88.66	1.4e+04	47506	AC018082 Drosophila melano
	+	65.00	79.63	4.5e+04	148439	AC094736 Rattus norvegicus	gb_htg:AC095730	+	64.50	87.76	1.6e+04	52562	AC095730 Rattus norvegicus



gb_ba:AE0000790	64.50	87.60	1.6e+04	53561	AE0000790 Borrelia burgdorferi	gb_htg:AC103547	64.00	83.88	2.6e+04	72205	AC103547 Rattus norvegicus	
gb_pr:AL162740	64.50	84.78	2.3e+04	73607	AL162740 Human DNA sequence	gb_pr:AC007022	-	-	-	72541	AC007022 Homo sapiens PAC	
gb_htg:AP004612	64.50	81.99	3.3e+04	100839	AP004612 Oryza sativa chromo	gb_htg:AC094243	64.00	83.34	2.8e+04	77057	AC094243 Rattus norvegicus	
gb_pr:AC019127	64.50	81.62	3.5e+04	105118	AC019127 Homo sapiens BAC cl	gb_htg:AC096701	-	64.00	83.15	2.9e+04	78480	AC096701 Rattus norvegicus
gb_ba:AP003366	64.50	81.25	3.6e+04	109589	AP003366 Staphylococcus aure	gb_htg:AC094818	-	64.00	82.71	3.0e+04	82432	AC094818 Rattus norvegicus
gb_ba:AP003338	64.50	80.67	3.9e+04	117016	AP003338 Staphylococcus aure	gb_htg:AC103561	+	64.00	82.27	3.2e+04	86628	AC103561 Rattus norvegicus
gb_htg:OSUN00099	64.50	80.30	4.1e+04	121992	AL606656 Oryza sativa chromo	gb_htg:AC098423	+	64.00	82.07	3.3e+04	88569	AC098423 Rattus norvegicus
gb_htg:AC020029	64.50	79.84	4.4e+04	128469	AC020029 Drosophila melanoga	gb_ro:AF332860S2	+	64.00	81.93	3.4e+04	89594	AF332861 Mus musculus diad
gb_htg:AC105719	64.50	79.62	4.5e+04	131728	AC105719 Rattus norvegicus c	gb_htg:AC108306	+	64.00	81.43	3.6e+04	95187	AC108306 Rattus norvegicus
gb_pr:HS203C2	64.50	78.43	5.2e+04	150620	274696 Human DNA sequence fr	gb_pl:ATT6H20	+	64.00	81.13	3.7e+04	98461	AL096859 Arabidopsis thalli
gb_htg:AC097260	64.50	78.24	5.3e+04	153876	AC097260 Homo sapiens chromo	gb_htg:AC091338	+	64.00	80.80	3.9e+04	102195	AC091338 Rattus norvegicu
gb_htg:AC023964	64.50	78.22	5.4e+04	154240	AC023964 Homo sapiens chromo	gb_htg:AC098618	+	64.00	80.33	4.1e+04	107782	AC098618 Rattus norvegicu
gb_pr:AC093901	64.50	77.83	5.6e+04	161153	AC093901 Homo sapiens BAC cl	gb_htg:AC105506	+	64.00	80.29	4.1e+04	108271	AC105506 Rattus norvegicu
gb_in:AC006170	64.50	77.44	5.9e+04	168471	AC006170 Drosophila melanoga	gb_htg:AC106364	-	64.00	80.08	4.2e+04	110875	AC106364 Rattus norvegicu
gb_htg:AC094714	64.50	77.31	6.0e+04	171046	AC094714 Rattus norvegicus c	gb_htg:AC010249	-	64.00	80.07	4.2e+04	111081	AC010249 Homo sapiens chr
gb_htg:AC097394	64.50	77.28	6.0e+04	172061	AC097394 Rattus norvegicus c	gb_htg:AC103074	-	64.00	79.85	4.4e+04	113766	AC103074 Rattus norvegicu
gb_in:AC009203	64.50	77.25	6.1e+04	172061	AC009203 Drosophila melanoga	gb_htg:AC095797	-	64.00	79.61	4.5e+04	116930	AC095797 Rattus norvegicu
gb_pr:AC017089	64.50	76.94	6.3e+04	178320	AC017089 Homo sapiens BAC cl	gb_htg:AC103495	+	64.00	79.53	4.5e+04	118002	AC103495 Rattus norvegicu
gb_htg:AC094933	64.50	76.84	6.4e+04	180191	AC094933 Rattus norvegicus c	gb_htg:AC096140	+	64.00	79.29	4.7e+04	121249	AC096140 Rattus norvegicu
gb_in:AC007817	64.50	76.83	6.4e+04	180499	AC007817 Drosophila melanoga	gb_htg:AC105708	+	64.00	79.04	4.8e+04	124695	AC105708 Rattus norvegicu
gb_htg:AC105841	64.50	76.29	6.9e+04	191747	AC105841 Rattus norvegicus c	gb_pr:HSJ685D6	+	64.00	78.98	4.9e+04	125487	AL049642 Human DNA sequen
gb_htg:AC021187	64.50	76.14	7.0e+04	195121	AC021187 Homo sapiens chromo	gb_pr:AC002386	+	64.00	78.57	5.1e+04	131458	AC002386 Human BAC clone
gb_pr:AC002524	64.50	75.81	7.3e+04	202540	AC002524 Homo sapiens x BAC	gb_htg:AC094240	+	64.00	78.51	5.2e+04	132359	AC094240 Rattus norvegicu
gb_htg:PFMAL13P6	64.50	75.72	7.4e+04	204652	AL049183 Plasmodium falcipar	gb_htg:AC098633	+	64.00	78.40	5.2e+04	132359	AC098633 Rattus norvegicu
gb_htg:AC106680	64.50	75.61	7.5e+04	207064	AC106680 Rattus norvegicus c	gb_htg:AP003368	+	64.00	78.33	5.3e+04	135057	AP003368 Oryza sativa chr
gb_htg:AC107077	64.50	75.61	7.5e+04	207197	AC107077 Homo sapiens chromo	gb_htg:AC098756	+	64.00	77.82	5.6e+04	143072	AC098756 Rattus norvegicu
gb_htg:AC095246	64.50	75.37	7.7e+04	221668	AC095246 Mus musculus chromo	gb_htg:AC106306	+	64.00	77.81	5.7e+04	143318	AC106306 Rattus norvegicu
gb_in:AE003761	64.50	75.07	8.0e+04	220116	AE003761 Drosophila melanoga	gb_htg:AC105735	-	64.00	77.66	5.8e+04	145680	AC105735 Rattus norvegicu
gb_htg:AC027284	64.50	74.89	8.2e+04	224614	AC027284 Mus musculus chromo	gb_htg:AC103147	+	64.00	77.57	5.8e+04	147237	AC103147 Rattus norvegicu
gb_in:AE003767	64.50	74.88	8.2e+04	224840	AC027284 Mus musculus chromo	gb_ba:D90916	+	64.00	77.51	5.9e+04	148142	D90916 Synchocystis sp.
gb_htg:AC098730	64.50	74.87	8.4e+04	225155	AE003767 Drosophila melanoga	gb_pr:AF274857	+	64.00	77.47	5.9e+04	148526	AF274857 Homo sapiens chr
gb_in:AC003660	64.50	74.69	8.4e+04	229788	AC098730 Mus musculus chromo	gb_htg:AC012240	+	64.00	77.49	5.9e+04	148804	AC012240 Homo sapiens clo
gb_ba:AJ414158	64.50	74.60	8.5e+04	231972	AE003660 Drosophila melanoga	gb_htg:AC106365	+	64.00	77.44	5.9e+04	149352	AC106365 Rattus norvegicu
gb_ba:AP0000984	64.50	74.49	8.6e+04	235050	AE003660 Drosophila melanoga	gb_htg:AC0979128	+	64.00	77.33	6.0e+04	151163	AC0979128 Oryza sativa chr
gb_vt:AF198100	64.50	73.83	9.4e+04	253050	AP000984 Sulfolobus tokodaii	gb_pl:AC079128	+	64.00	77.31	6.0e+04	151589	AC079128 Oryza sativa chr
gb_pr:AF196295	64.50	72.67	1.1e+05	288539	AF198100 Fowlpox virus, comp	gb_htg:AC079022	+	64.00	77.25	6.1e+04	152680	AC079022 Oryza sativa chr
gb_pl:FMU981120	64.50	71.32	1.3e+05	335913	AF196295 Sequence 2 from Pat	gb_htg:AC106397	+	64.00	77.20	6.1e+04	153540	AC106397 Rattus norvegicu
gb_pat:AX196296	64.50	71.32	1.3e+05	335913	AX196296 Sequence 3 from Pat	gb_htg:AC106212	+	64.00	77.13	6.2e+04	154715	AC106212 Rattus norvegicu
gb_pl:FMU88110	64.00	128.29	87.90	482	U98110 Fusarium merismoides var	gb_htg:AC097562	+	64.00	77.06	6.2e+04	155909	AC097562 Rattus norvegicu
gb_ba:AF12696	64.00	126.66	108.28	579	AC031736 Sequence 19 from pater	gb_htg:AC097562	+	64.00	77.04	6.2e+04	156334	AC097002 Rattus norvegicu
gb_in:AY070642	64.00	125.54	125.01	657	AC031736 Sequence 3 from pater	gb_htg:AC061114	+	64.00	76.99	6.3e+04	157088	AC106114 Rattus norvegicu
gb_ba:AB047188	64.00	119.95	256.18	1235	AY070642 Drosophila melanogast	gb_htg:AC097619	+	64.00	76.99	6.3e+04	157179	AC097619 Rattus norvegicu
gb_ba:AB047191	64.00	118.70	300.72	1422	AB047188 Marine CFB-group bact	gb_htg:AC0974218	-	64.00	76.88	6.4e+04	159185	AC074218 Mus musculus chr
gb_pl:YSCC9TH1	64.00	118.70	300.72	1422	AB047191 Marine CFB-group bact	gb_htg:AC098073	+	64.00	76.88	6.4e+04	160862	AC098073 Rattus norvegicu
gb_to:AF121893	64.00	118.16	321.97	1510	L42133 Saccharomyces cerevisia	gb_htg:AC099162	+	64.00	76.71	6.5e+04	162127	AC099162 Rattus norvegicu
gb_ba:AF455356	64.00	116.05	422.32	1917	AF121893 Rattus norvegicus seq	gb_pr:AC026799	+	64.00	76.67	6.5e+04	162918	AC026799 Homo sapiens chr
gb_ba:AF155013	64.00	113.87	558.43	2451	AF455356 Alicaligenes faecalis	gb_htg:AC105675	+	64.00	76.62	6.6e+04	163846	AC105675 Rattus norvegicu
gb_ba:AF190145	64.00	113.40	593.01	2584	AF195013 Helicobacter pylori s	gb_htg:AC099136	+	64.00	76.58	6.6e+04	164675	AC099136 Rattus norvegicu
gb_ba:AF109145	64.00	113.30	601.10	2615	AF109145 Vibrio cholerae 569B	gb_htg:AC102961	+	64.00	76.57	6.6e+04	164714	AC102961 Rattus norvegicu
gb_pr:BC009895	64.00	112.57	660.24	2840	BC009895 Homo sapiens, phospho	gb_htg:AC099146	+	64.00	76.56	6.6e+04	165055	AC099146 Rattus norvegicu
gb_ov:AB055658	64.00	112.39	675.59	2898	AB055658 Cyprinus carpio GRK7	gb_htg:AC097035	+	64.00	76.42	6.7e+04	167579	AC097035 Rattus norvegicu
gb_ov:AB046568	64.00	111.90	719.22	3062	L42546 Xenopus laevis LIM clas	gb_htg:AC094724	-	64.00	76.32	6.8e+04	169452	AC094724 Rattus norvegicu
gb_ba:AB046568	64.00	108.83	1.1e+03	4330	AB046568 Hydrogenobacter therm	gb_htg:AC094122	+	64.00	76.32	6.8e+04	169552	AC094122 Rattus norvegicu
gb_vt:CVU10029	64.00	108.64	1.1e+03	4420	U10029 Chlorella virus PBCV-1	gb_htg:AC099185	+	64.00	76.32	6.8e+04	169565	AC099185 Rattus norvegicu
gb_pl:AF267125	64.00	107.94	1.2e+03	4783	AF267125 Candida albicans pota	gb_htg:AC094951	+	64.00	76.29	6.9e+04	170000	AC094951 Rattus norvegicu
gb_pat:AR120577	64.00	105.62	1.6e+03	6216	AR120577 Sequence 1 from pater	gb_htg:AC106360	-	64.00	76.28	6.9e+04	170338	AC106360 Rattus norvegicu
gb_pat:AX0719922	64.00	104.64	1.8e+03	6942	AF099222 Sequence 10 from Pate	gb_htg:AC073047	+	64.00	76.09	7.0e+04	173854	AC073047 Homo sapiens chr
gb_ba:AE004243	64.00	101.23	2.8e+03	10208	AE004243 Vibrio cholerae chro	gb_htg:AC102971	+	64.00	76.07	7.1e+04	174336	AC102971 Rattus norvegicu
gb_ba:AE007784	64.00	100.93	2.9e+03	10553	AE007784 Clostridium acetobut	gb_htg:AC096983	+	64.00	76.05	7.1e+04	174664	AC096983 Rattus norvegicu
gb_ba:AE006197	64.00	100.85	3.0e+03	10655	AE006197 Pasteurella multocid	gb_pr:AC024359	-	64.00	75.92	7.2e+04	177303	AC024359 Homo sapiens chr
gb_ba:AE000566	64.00	99.63	3.5e+03	12216	AE000566 Helicobacter pylori	gb_htg:AC094510	-	64.00	75.68	7.2e+04	182146	AC094510 Rattus norvegicu
gb_ba:U32710	64.00	99.24	3.6e+03	12771	AE002526 Neisseria meningitid	gb_htg:AC095656	+	64.00	75.66	7.4e+04	182685	AC095656 Rattus norvegicu
gb_ba:U32710	64.00	98.54	4.0e+03	13818	U32710 Haemophilus influenzae	gb_htg:AC106074	+	64.00	75.64	7.4e+04	183095	AC106074 Rattus norvegicu
gb_ba:AF282852	64.00	95.94	5.6e+03	18530	AE001481 Helicobacter pylori,	gb_htg:AC013555	+	64.00	75.62	7.5e+04	183251	AC013555 Homo sapiens chr
gb_ba:AF282852	64.00	94.69	6.5e+03	21347	AF282852 Helicobacter pylori,	gb_ro:AC087416	+	64.00	75.62	7.5e+04	183347	AC087416 Mus musculus clo
gb_ba:AC000108	64.00	89.44	1.3e+04	38570	AC000108 Helicobacter pylori	gb_pr:AC093515	-	64.00	75.58			







```

gb_in:AE003834 + 63.50 71.55 1.2e+05 257324 ! AE003834 Drosophila melanoga
gb_in:AE003573 - 63.50 71.39 1.3e+05 262059 ! AE003573 Drosophila melanoga
gb_in:AE003839 - 63.50 71.39 1.3e+05 262228 ! AE003839 Drosophila melanoga
gb_htg:PFWAL13P4 + 63.50 70.39 1.4e+05 293431 ! AL049181 Plasmodium falcipar
gb_htg:PFWAL13P4 - 63.50 70.39 1.4e+05 293431 ! AL049181 Plasmodium falcipar
gb_ba:AP002551 - 63.50 70.32 1.5e+05 295741 ! AP002551 Escherichia coli O1
gb_in:AE003602 + 63.50 70.26 1.5e+05 297832 ! AE003602 Drosophila melanoga
gb_in:AE003606 - 63.50 70.19 1.5e+05 300265 ! AE003606 Drosophila melanoga
gb_ba:AP003133 - 63.50 70.08 1.5e+05 303750 ! AP003133 Staphylococcus aure
gb_in:AE003781 - 63.50 69.85 1.5e+05 311800 ! AE003781 Drosophila melanoga
gb_ba:CJ11168X1 + 63.50 69.78 1.6e+05 314150 ! AL139074 Campylobacter jejuni

seq_name: gb_ba:ECOETOXHL
seq_documentation_block:
LOCUS ECOETOXHL 1262 bp DNA linear BCT 18-NOV-1994
DEFINITION Escherichia coli (strain 41) heat-labile enterotoxin type IIB
(LT-IIB) A and B chain genes, complete cds.
ACCESSION M28523
VERSION M28523.1 GI:576584
KEYWORDS LT-IIB gene; enterotoxin type IIB.
SOURCE Escherichia coli (strain 41) DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 1262)
AUTHORS Pickett,C.L., Twiddy,E.M., Coker,C. and Holmes,R.K.
TITLE Cloning, nucleotide sequence, and hybridization studies of the type
IIB heat-labile enterotoxin gene of Escherichia coli
JOURNAL J. Bacteriol. 171 (9), 4945-4952 (1989)
MEDLINE 89359131
COMMENT On Nov 28, 1994 this sequence version replaced gi:341953.
FEATURES
Location/Qualifiers
1..1262
/organism="Escherichia coli"
/strain="41"
/db_xref="taxon:562"
/clone="pcp4185"
1..12
/gene="LT-IIB"
1..1262
/gene="LT-IIB"
26..31
/gene="LT-IIB"
113..904
/gene="LT-IIB"
/note="A chain of heat-labile enterotoxin type IIB"
/codon_start=1
/transl_table=11
/protein_id="AAA53285.1"
/db_xref="GI:576585"
/translation="MAKVISFFISLFLISFPLIYANDYFRADSRTPDEVRRSGLLIPRG
QDEYRGPINLYDTHARGATNRYNDGYSTTTTLRQHLHGNMUGLYNEYI
IYVAAALFDVNGVLGRYPSPSENYAALGGIPLSQIIGWYRVSFGEIABGGMHRN
RYRDLFLGLAANEDGYRAGPDGFPAWEVPWFREFAPNSLPNNKASSDTTCA
SLTNKLSQHDLDLDFKRYIKRKTTLTLISNNDGFFSNNGKDEL"
113..172
/gene="LT-IIB"
173..901
/gene="LT-IIB"
/note="A chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"
894..962
/gene="LT-IIB"
894..1262
/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/codon_start=1
/transl_table=11
/product="enterotoxin"
/protein_id="AAA53286.1"
/db_xref="GI:576586"

```

```

/translation="MSFKKIIFKAFVIMAAVLSVQAHAGASOFFKDNCRRTASLVEGV
ELXYISDINNNTDGMVYVSTGGWRISRAKDYPDNVMTAEMRKIAMAAVLSGRVYN
MCASPSPNPVIAIELEAE"
mat_peptide
963..1259
/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"
BASE COUNT 369 a 195 c 285 g 413 t
ORIGIN
alignment_scores:
Quality: 181.00 Length: 141
Ratio: 2.662 Gaps: 11
Percent Similarity: 48.227 Percent Identity: 43.972
alignment_block:
US-09-528-682-4 x ECOETOXHL ..
Align seg 1/1 to: ECOETOXHL from: 1 to: 1262
34 GlyArgTyrSerAsnTyr.....AlaLeuIleSerA 44
|||||.....|
470 GGCAGGTATAGTCCGTATCCAGTGAAATGATATGCTGCATTAGGTGG 519
40 .....AlaLeuIleSerA 44
520 TATTCCTCTCACAATAATTCGGTGGTATACAGTATCTTCGCGTGCTA 569
44 laGluGlyGlyMetAsp.....ArgAspLeuPheGly 54
|||||.....|
570 TAGAGGGGGGATGCATCGAACAACAGGATTACAGAAGAGATTTATTAGA 619
55 SerAlaAsnIle.....
620 GGTTATCTGCTCTCCTAATGAGGATGGCTATCGAATTGCTGGATTTC 669
59 AspGlyPhePro.....GluVal.....ArgGluPhe..... 67
|||||.....|
670 CGACGGATTTCCTCGCGTGGGAAGAGTTCCTCGAGAGAAATTCGCCCTA 719
68 ..AsnSerLeuProAsnAsnLysAlaSerSerAspThr.....AlaSer 81
::: |||||.....|
720 ACTCTTGTCTACCAATAATAAAGCTTCACAGTGATACCTGTGCTCT 769
82 Leu...AsnLys.....GlnHisAsp...AlaAspPheLysLysTyrI 94
||| |||||.....|
770 TTACAAATAAATATCATCACAGCATGATTAGCTGATTTAAGAAATATAT 819
94 eLys.....LeuLeuIleAsnAsnAspGlyP 103
||| |||||.....|
820 AAAGAGAAAATTTACCTTGATGACTTTACTAAGTATTAATAATGATGTT 869
103 hePheSerAsnAsnGlyGlyLys 110
|||||.....|
870 TTTTTCAAATAATGGAGGAAAA 892
seq_name: gb_ba:AE0000850
seq_documentation_block:
LOCUS AE0000850 11069 bp DNA linear BCT 15-NOV-1997
DEFINITION Methanobacterium thermoautotrophicum from bases 631143 to 642211
(section 56 of 148) of the complete genome.
ACCESSION AE0000850 AE0000666
VERSION AE0000850.1 GI:2621794
KEYWORDS
SOURCE Methanothermobacter thermautotrophicus
ORGANISM Methanothermobacter thermautotrophicus
Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
Methanobacteriaceae; Methanothermobacter.
REFERENCE 1 (bases 1 to 11069)
AUTHORS Smith,D.R., Doucette-Stamm,L.A., Deloughery,C., Lee,H., Dubois,J.,
Aldredge,T., Bashirzadeh,R., Blakely,D., Cook,R., Gilbert,K.,

```



Harrison,D., Hoang,L., Keagle,P., Lumm,W., Pothier,B., Qiu,D., Spadafora,R., Vicaire,R., Wang,Y., Wierzbowski,J., Gibson,R., Jivani,N., Caruso,A., Bush,D., Reeve,J.N. et al.  
Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics  
J. Bacteriol. 179 (22), 7135-7155 (1997)  
98037514  
2 (bases 1 to 11069)  
Smith,D.R.  
Direct Submission  
Submitted (10-AUG-1997) Genomics and Technology Development, Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154-8448, USA  
FEATURES  
source  
1..11069  
/organism="Methanothermobacter thermoautotrophicus"  
/strain="delta H"  
/db\_xref="taxon:145262"  
/clone="MTH"  
98..1009  
/gene="MTH708"  
98..1009  
/gene="MTH708"  
/note="Function Code:4.02 - Nucleotide Metabolism, Pyrimidine metabolism ; similar to,  
gp\_new:GI:c313024:g1945648 LN:BSZ94043, p()=3.8E-57, pid=428"  
/codon\_start=1  
/transl\_table=11  
/product="thioredoxin reductase"  
/protein\_id="AAB85213.1"  
/db\_xref="GI:2621795"  
/translation="MWTYDMIVIGAGPAGLTAGTGYGRCSSVLMLDKGPAGGLGLE  
VPMENYPCFEMAGSLVTRMKQATVAELREMEVEKEKGDVFTVYKSRDYFTA  
SAIFATGSKRHQGVGNDLLRGVCYCACDGPPLYKGRKVLVGGVGNAAQAVF  
LKNIGCDYSIVHRRDELRAKYLDQKLEMEIPVITWSVVKVEIGDGERVEEVIHNRV  
TGRDELVDGVFTAIERGEPLNQLADVLGVEVDKGYIITDKFORTNPLVVAAGDIT  
GGLNQWTFACAGGAATAYAYREIQSY"  
1313..1873  
/gene="MTH709"  
1313..1873  
/gene="MTH709"  
/note="Function Code:4.01 - Nucleotide Metabolism, Purine metabolism ; similar to, pir:LN:F64496 AC:F64496, p()=1.8E-52, pid=498"  
/codon\_start=1  
/transl\_table=11  
/product="GMP synthetase, subunit A"  
/protein\_id="AAB85214.1"  
/db\_xref="GI:2621796"  
/translation="MFMLIINHGQYNHRTUHYLIQIPSELVPNTPTLDEIISRDL  
PEGLLGGPSLERSGNCVEYIQLEIPILGICLGHQLIARAYGEVATAEASVQI  
EILEIDEDIFRGLGPRMSWASHKDEYKLPDEFDLATSSICEVAMKHKDKPVYG  
IQFHEVHTREGRHVFENFVDVCRS"  
1976..2911  
/gene="MTH710"  
1976..2911  
/gene="MTH710"  
/note="Function Code:4.01 - Nucleotide Metabolism, Purine metabolism ; similar to, pir:LN:B64441 AC:B64441, p()=1.3E-109, pid=648"  
/codon\_start=1  
/transl\_table=11  
/product="GMP synthetase, subunit B"  
/protein\_id="AAB85215.1"  
/db\_xref="GI:2621797"  
/translation="MMLMNPSETEEAEEIRSVGNKEATIALSGGVDSVASVLA  
GRAIGNLTAFVDHGLLREGEAEYVRFTSERLNFRIIDASEEFLKELEGYDPEEK  
RKIIQGVTEVFVEAEYRGLVQGTIADPIWIESEGIKSHHNVLPGLHVLVEIE  
PIRELYKDEVEYIGLEGLPMTQIQPPGGLAVRIVGETREKIEICRRANAIVE  
BEVRSGLHESLMQYFAVLTDTMTQVKGQVDRDFGLVYVVRMVESILDATASVPELWP  
DLIKRISKRIETAEIPEVTHVALSDKPKPSTIEFA"  
3078..3533

CDS  
/gene="MTH711"  
3078..3533  
/gene="MTH711"  
/note="Function Code:10.02 - Metabolism of Macromolecules, transcription--mRNA synthesis and modification (includes regulators) ; similar to, gp:GI:e266415:g1550715, p()=5.8E-38, pid=498"  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional regulator"  
/protein\_id="AAB85216.1"  
/db\_xref="GI:2621798"  
/translation="MKDRIPIFSVAKNRVEMVERIELSDDEWREILDPEAFVRKARR  
TEPPTGKYHDLHDGIYRCICCGTDLFDSETKFDGSGWPSFYDVVSEHNKLRDR  
SLGMYRCVLCARCDALHGHVDDGPRPTGKRYCMNSAALKEIPRQIG"  
3535..3786  
/gene="MTH712"  
3535..3786  
/gene="MTH712"  
/note="Function Code:14.00 - Unknown, ; similar to,  
gp:GI:g1652640 LN:D90907, p()=0.074, pid=208"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAB85217.1"  
/db\_xref="GI:2621799"  
/translation="MQEGAGMAEDFGRYWVAGLIMAVATEFINTLYIIGTWPRNFPY  
EWILGLLGIAGIAMFILYPMIYGRLEFVYIHFISGEG"  
complement(3787..4149)  
/gene="MTH713"  
complement(3787..4149)  
/gene="MTH713"  
/note="Function Code:14.00 - Unknown, ; similar to,  
gp:GI:g1695706 LN:D88814, p()=0.97, pid=258"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAB85218.1"  
/db\_xref="GI:2621800"  
/translation="MYSRMQRHQVORIGYNVRFRGSLILLRGGDRPQBLRDEFMDIE  
RSLTRVFRDISIEIPDGLADLHLIFKAVGTYLEAREFLFYDGGDDHVFHSGLM  
INEANELMNQAAAFNGL"  
complement(3799..3875)  
/note="77 bp direct repeat includes part of MTH713  
(unknown) ; 95% ID to interval 649558-649482"  
complement(3913..4012)  
/note="100 bp direct repeat includes part of MTH713  
(unknown) ; 81% ID to interval 649675-649573"  
complement(4563..9926)  
/gene="MTH714"  
complement(4563..9926)  
/gene="MTH714"  
/note="Function Code:9.10 - Metabolism of Cofactors and Vitamins, Porphyrin and chlorophyll metabolism ; similar to, gp:GI:g1652115 LN:D90902, p()=1.2E-73, pid=168"  
/codon\_start=1  
/transl\_table=11  
/product="magnesium chelatase subunit"  
/protein\_id="AAB85219.1"  
/db\_xref="GI:2621801"  
/translation="MTYNVTLVGNAGNODTSGVRVDFVNGVRVGSYTIPLRSGET  
TIVVFWPTVGSHTLQIILDPENHIOETNETNRRILLNTVNOALPDLICKSTELV  
ADLIYNTFTTARISNGIPANAGVRVDFVNGVRVGSYTIPLRSGETTIVVFNWT  
PMTVGSHTLQIILDPENHIOETNETNRRILLNTVNOALPDLICKSTELV  
VHDANVLYGAPVRKQIQIRSEQLTKMSDNELLQVSACDIFLGEYLSDSASRLQ  
KLLSNPGTITKGGTIFLLEPAIVNPAYTSLMRYSNIKGEYILSNIDELLIDYQ  
NTKKANFEHVLNVISSNTIGIFSLEYNMAVLYKDINNNAFVQIVMALNLIGVFN  
TGYYRPPFWATNSGEPYGIYRWGNVLEYNKYFRANATGIGLESKYVDNOKLO  
PYIADILEARGFNVPVAYATPEQLKVNVESTDPADVESFANPSKKYVYDA  
IVSMESGGIDFTNVTKFFELINVPVKALHSYLLAEQWDLSTLGRQLTGDRWM  
HIVASESGQIIDTNTSEPTDPTGLSYSLYEIPANIEHLADEIKGWKLRRL  
SNNEKKVALIYYNPPGKNIGSSYLDVITSYINLLNKLKNGYTVENIPANTTILHD  
MTIQGINVAPWAPGELEKLNAILYPVDKYMWFSLKLEITIKIQVIEGPGYIGEL



















ORGANISM  
Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE  
1 (bases 1 to 13705)  
AUTHORS  
Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,  
Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,  
Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,  
Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,  
Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,  
Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,  
O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,  
Saunders,D., Showkhen,R., Smaldon,N., Smith,A., Sonhammer,E.,  
Staden,R., Sulston,J., Thierrey-Mieg,J., Thomas,K., Vaudin,M.,  
Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,  
Wilkinson-Sproat,J. and Wohldman,P.  
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans  
TITLE  
Nature 368 (6466), 32-38 (1994)  
JOURNAL  
94150718  
MEDLINE  
2 (bases 1 to 13705)  
REFERENCE  
Bradshaw,H., Becker,M. and Fronick,W.  
AUTHORS  
The sequence of C. elegans cosmid M4  
TITLE  
Unpublished (1998)  
JOURNAL  
3 (bases 1 to 13705)  
REFERENCE  
Waterston,R.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (22-OCT-1998) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT  
Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
e-mail: rwnematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we provide a small overlap between  
neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate  
chemistry; an attempt was made to resolve all sequencing problems,  
such as compressions and repeats; all regions were covered by  
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' clone is Y67D8C;3' clone is Y67D8B. Actual start of this  
clone is at base position 1 of CELM4; actual end is at 13705 of  
CELM4

The tandem repeat region from 2818 to 6039 of this sequence is  
deleted. Digest data with XhoI of the overlying YAC Y67D8 showed  
the tandem region to be approximately 18.6 kilobases indicating a  
deletion in M4 of approximately 15.4 kb. The tandem contains single  
stranded regions.

NOTES:

Coding sequences below are predicted from computer analysis, using  
the program Genefinder(P. Green and L. Hillier, ms in preparation).  
Location/Qualifiers  
1..13705  
/organism="Caenorhabditis elegans"  
/strain="Bristol N2"  
/db\_xref="taxon:6239"  
/chromosome="IV"











```

15579 CATAAAAAGCATCAACACAACTCTCAAAGGTAATGTAATAATGCAAAAAGC 15530
29 .....GluValAlaLeuAspGlyArgTyrSerAsn 38
15529 TCCTAGCCCAAAACATCCAGGAATCCAGGACACAAATGAGAAGTTCAAAT 15480
39 TyrAlaLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheGlyse 55
15479 CTAAAGGATATAGGTATGAAGAGAGATAAGATTCTCAACTTAAGGGTTC 15430
55 rAlaAsnIle.....AspGlyPheProGluValArg 66
15429 AGTAATAATCTTCAACAAAAATTATAGAAGAAAATTTCCCTAACTTAAAGA 15380
66 lUpheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeu 82
15379 AGAGAAGTCCATAAACAACACAGAGAGCTACAGAACTCCAATAGATTG 15330
83 Asn...LysGlnHisAspAlaAspPheLysLys 92
15329 GACTAGAAGAAATTCCTCTGTGATTACAGAAA 15297

```

seq\_name: gb\_pat:AX067440

```

seq_documentation_block:
LOCUS AX067440 28626 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 15 from Patent WO0078968.
ACCESSION AX067440
VERSION AX067440.1 GI:12545060
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE
AUTHORS Lagace,R.E., Patterson,C. and Berg,K.L.
TITLE Nucleotide sequences of moraxella catarrhalis genome
JOURNAL Patent: WO 0078968-A 15 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 28626
/organism="Moraxella catarrhalis"
/db_xref="taxon:480"
BASE COUNT 8524 a 5568 c 6119 g 8415 t
ORIGIN

```

```

alignment_scores:
Quality: 78.00 Length: 116
Ratio: 1.300 Gaps: 6
Percent Similarity: 51.724 Percent Identity: 29.310

```

```

alignment_block:
US-09-528-682-4 x AX067440 ..
Align seg 1/1 to: AX067440 from: 1 to: 28626
5 ValArgIleGlnAspAlaTyrGluProIleAla.....AsnThrAs 18
8394 GTTAGCAGCTCGTGATGACCTAAATATGCTGCTAAATTTTCGTCAAAACCA 8443
18 nThrThrThrGlnPheLeuAsnMetGly.....AsnGluValAlaLeuA 33
8444 CCAACAACACCCCTGTTTGTGTCATGGGTTATTTAAACCCCTGTTGAAATCA 8493
33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAsp 49
8494 TTGGT...TATGATATTTTGCCTAATTTGTGCTGATGCTGCGGTGAT 8540
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG1 66
8541 GCANTTTGTATG.....GTGGAT..... 8558

```

```

66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83
8559 .....TTACCGCCCAACAGCAGATGGCTATGCAAAATCACTGA 8598
83 snLysGlnHisAspAlaAspPheLysLysTyrIle..... 94
8599 CCGAGCGTACCGATCATCCCATGAATCAAAATTTTTCGTGGCACCACCA 8648
95 .....LysLeuLeuIleAsnAsnAspGlyPhe 103
8649 ACAAAATCAGATCGCGTACCAAGGTTATCGAAACATGTAGTGGGTTT 8696
seq_name: gb_htg:CNS07EFS

```

```

seq_documentation_block:
LOCUS CNS07EFS 155031 bp DNA linear HTG 06-FEB-2001
DEFINITION Oryza sativa chromosome 12 clone OSJNBa0052H10, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AL513402
VERSION AL513402.1 GI:12711300
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriartoideae; Oryzaceae; Oryza.
1 (bases 1 to 155031)
Salse,J., Choise,N., Orjeda,G., Regad,F., Lorieux,M., Cooke,R.,
Delseny,M., Robert,C., Brottier,P., Wincker,P., Cruaud,C.,
Artiguenave,F., Saurin,W., Salanoubat,M., Quetier,F. and
Weissenbach,J.
Oryza sativa chromosome 12 sequencing

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

```

```

COMMENT
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
Contigs composition :
122605 bp contig from 1 to 122605
31426 bp contig from 123606 to 155031.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1. 155031
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OSJNBa0052H10"
/clone_lib="CUGI Nipponbare BAC"
BASE COUNT 43146 a 34070 c 33654 g 43146 t 1015 others
ORIGIN

```

```

alignment_scores:
Quality: 77.00 Length: 111
Ratio: 1.481 Gaps: 5
Percent Similarity: 46.847 Percent Identity: 28.829

```

```

alignment_block:
US-09-528-682-4 x CNS07EFS/rev ..

```



Align seg 1/1 to reverse of: CNS07EFS from: 1 to: 155031

46 GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePr 62

8274 GGTTCCTACCTAGCAATCTTGGTGGTGACAAAGGTACACCGTGCTCC 8225

62 ogluValArgGluPhe...AsnSerLeuProAsnAsnLysAlaSerSerA 78

8224 AGAGTTGTAGAGTTCTATCAAGTCTCATGACGCTGAAGCCAAAGG 8175

78 spThrAlaSerLeu.AsnLysGlnHis..... 86

8174 ACACAACCTCTCTGGGATCAACAACATCAAGTGTCTGTGATGAGAAGC 8125

86 ..... 86

8124 AACATGATGGAGAGATTGAGATAGATACATTCCTCTTAGCTGTAAG 8075

87 .....AspAlaAspPheLysLysTyrile..... 94

8074 TTATTCAGATGATGCTACCTTCAGAAATATATACTACTGAATTCCTTT 8025

95 ...LysLeuLeuIleAsnAsnAsp..... 101

8024 AGTCCACTGCTAATTCAGGTAGATTATTATGAGGACTTACTTTTATATA 7975

102 .....GlyPhePheSerAsnAsnGlyGly 109

7974 AAAAAATGGATTCTTTTCAGGTCAAGTGGGA 7944

seq\_name: gb\_in:TAY15795

seq\_documentation\_block:

LOCUS TAY15795 957 bp DNA linear INV 11-MAY-2001

DEFINITION Theileria annulata spm2 gene, partial.

ACCESSION Y15795

VERSION Y15795.1 GI:2673881

KEYWORDS spm2 gene; spm2 protein.

SOURCE Theileria annulata.

ORGANISM Theileria annulata.

REFERENCE 1 (bases 1 to 957)

AUTHORS Knight,P.A., Williamson,S.W., Brown,C.G., Bell-Sakyl,L., Kirvar,E.,

Kinnaird,J., Hussain,K., Hall,R., Boulter,N. and Tait,A.

TITLE Identification of a Theileria annulata antigen expressed in

multiple stages of the parasite life cycle

JOURNAL Exp. Parasitol. 90 (1), 110-121 (1998)

MEDLINE 98376467

REFERENCE 2 (bases 1 to 957)

AUTHORS Knight,P.A.

TITLE Direct Submission

Submitted (05-DEC-1997) P.A. Knight, Edinburgh University,

Department of veterinary clinical studies, Veterinary field

station, Easter Bush Roslin, Midlothian EH25 9RG, UK

Location/Qualifiers

1. .957

/organism="Theileria annulata"

/isolates="Ankara D/ macrochizont-infected leucocyte cell

line"

/specific\_host="cow, buffalo"

/db\_xref="taxon:5874"

/chromosomes="2"

/clone="lambda gt11-KP8"

/sub\_clone="KP8 insert subcloned in pGEM7ZF-SK+ (Promega)"

/clone\_lib="lambda gt11"

/dev\_stage="merozoite (in vitro derived)"

1. .957

/gene="spm2"

<1.>957

/gene="spm2"

/codon\_start=1

gene

CDS

/product="spm2 protein"  
/protein\_id="CAA75787.1"  
/db\_xref="GI:2673882"  
/db\_xref="SPTREMBL:O43950"  
/translation="RDSAKKYSDETHKESKDHREYKYNKNDNSKDYCIDSEAIK  
AVBEKAVIEAFDKLSEIKIGKEETCLDSVNREYTFDKNDTTLSTYQSOSQLYSRDIQ  
RLTNDQNNQSTDMNRYGKNDFKENNDTHGLKPFKDFSNVREYSKDKTVESNN  
PEFKGIISSTAYANTQENPYVTDHMSQNVYEISYENKTFDEFDSFNFQFNKGSQ  
FEYNFDGVDLFSNRDQDLSLPDDEQFNSLNNNVAVRDDDFAVGEMYDKEQDDNFSEQ  
LTNQDDKSELSCSTFEQNDLDDKNFDQTYTLGNCNFANFTRA"

BASE COUNT 397 a 150 c 180 g 230 t

ORIGIN

alignment\_scores:

Quality: 76.50 Length: 103

Ratio: 1.297 Gaps: 5

Percent Similarity: 57.282 Percent Identity: 26.214

alignment\_block:

US-09-528-682-4 x TAY15795 ..

Align seg 1/1 to: TAY15795 from: 1 to: 957

8 GlnAspAlaTyrGluProIleAlaAsnThrAsnThrThrThrGlnPhe.. 23

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

565 CAGAATGTGTATGAGGAGATTTTCATATGAGATAAAACTGATGAATTGA 614

24 .....LeuAsnMetGlyAsnGluValAlaLeuA 33

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

615 TAGTGAATCGTTTAAATTTTCAATTCAACAGGGGGAGCCAA.....TTTG 658

33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAsp 49

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

659 AATACAAATTTTATGAGTGTGTTTCAATAGACACCAACAGAT 708

50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgI 66

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

709 ...AGTTTATTC.....TACCAGATGAAGAACA 734

66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

735 ATTCAATAGTTTGAATAATAATAGTCTGTAAGAGATGACTTCGCCGTGG 784

83 snLysGlnHisAspAlaAspPheLys...LysTyrIleLysLeuLeuIle 98

||||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||

785 GAGAAATGTATGATAGGACACACAGATAACTTCGGACAGATTGACC 834

99 AsnAsnAsp 101

||||:||||

835 AACCAAGAT 843

seq\_name: gb\_ba:AF160726

seq\_documentation\_block:

LOCUS AF160726

DEFINITION

peanut witches'-broom phytoplasma heat shock protein GrpE (grpe),

heat shock protein DnaK (dnaK), and heat shock protein DnaJ (dnaJ)

genes, complete cds; and unknown gene.

ACCESSION AF160726

VERSION AF160726.1 GI:9621759

KEYWORDS

SOURCE

ORGANISM

peanut witches'-broom phytoplasma.

Bacteria; Firmicutes; Bacillus/Clostridium group: Mollicutes;

Acholeplasmatales; Acholeplasmataceae; Phytoplasma; 16SrII (Peanut

WB group).

REFERENCE 1 (bases 1 to 4974)

AUTHORS Chu,P.W. and Lin,C.P.

TITLE Cloning and Analysis of dnaK and dnaJ Genes of Phytoplasma

Associated with Peanut Witches' Broom

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4974)











BASE COUNT 19550 a 14643 c 14414 g 18723 t 5399 others  
ORIGIN

/db\_xref="taxon:10116"  
/clone="CH230-11A23"

alignment\_scores:  
Quality: 76.50 Length: 155  
Ratio: 1.109 Gaps: 5  
Percent Similarity: 44.516 Percent Identity: 21.935

alignment\_block:

US-09-528-682-4 x AC097767 ..

Align seg 1/1 to: AC097767 from: 1 to: 72729

3 PheThrValArgIleGlnAspLafYrGluProIleAlaAsnThrAsnTh 19  
||||| |||||||: ||| |||:|||||:  
7490 TTCACCAACAGATACAGAGATGGAAGAGATCTCAGAGCAGAAGA 7539

19 rThrThrGlnPheLeuAsnMetGly..... 27

7540 TTCATAGAAATCATGTCACCTCACTCAAGATATGTAAGCGGAAAA 7589

28 .....AsnGluValAlaLeuAspGlyArgTyrSer 37

7590 AGCTACTGTCCTCAACATACAGGAATCCAGGACTCAATGAGAGATCA 7639

38 AsnThrAlaLeuIleSerIaGluGlyMetAspArgAspLeuPheG 54

7640 AACCTAAGGATAATAGGTATAGAGAGAGTGAAGACTCCAGCTCAAAG 7689

54 ySerAlaAsnIle.....AspGlyPheProGluVala 65

7690 ATCATGTAATATCTTCAACAAAATCATAGAGAAAATTCCTCAACTAA 7739

65 rgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSer 81

7740 AGAA.....GAGATGCCAATAAAAAACAGAACCTACAGACTAGA 7783

82 LeuAsnLys.....GlnHis..... 86

7784 ACAATAAATATGGACCAAGAAAAACCAACATGTCAATAATAGTCAAA 7833

86 ..... 86

7834 ACATAAGTGCACAAAACAGAAAAAATACTAAAGCAGTAGAGGGAAA 7883

87 .....AspAlaAspPheLysLysTyrIleLysLeuLeu 97

7884 TGGCCCAAGTAACCTGTAAAGGCAGACCTACCAAAATTACACCAAACTCTC 7933

98 IleAsnAsnAspGly 102

7934 ACCAGAGACAAATGGA 7948

seq\_name: gb\_vl:AF410153

seq\_documentation\_block:

LOCUS AF410153 146454 bp DNA linear VRL 31-JAN-2002  
DEFINITION Swinepox virus isolate 17077-99, complete genome.  
ACCESSION AF410153  
VERSION AF410153.1 GI:18448493

KEYWORDS Swinepox virus.

SOURCE Swinepox virus

ORGANISM Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

Sulfoxvirus.

1 (bases 1 to 146454)

REFERENCE Afonso,C.L., Tulman,E.R., Lu,Z., Zsak,L., Osorio,F.A., Balinsky,C.,

AUTHORS Kutish,G.F. and Rock,D.L.

TITLE The genome of swinepox virus

J. Virol. 76 (2), 783-790 (2002)

MEDLINE

PUBMED 21624277

REFERENCE 11752168

AUTHORS

TITLE

JOURNAL

FEATURES

source

21624277  
11752168  
2 (bases 1 to 146454)  
Afonso,C.L., Tulman,E.R., Lu,Z., Balinsky,C., Osorio,F.A., Zsak,L.,  
Kutish,G.F. and Rock,D.L.  
Direct Submission  
Submitted (09-AUG-2001) African Swine Fever Research, Plum Island  
Animal Disease Center, U.S. Dept. Agriculture, Agricultural  
Research Service, P.O. Box 848, Greenport, NY 11944-0848, USA  
Location/Qualifiers  
1. .146454  
/organism="Swinepox virus"  
/isolate="17077-99"  
/db\_xref="taxon:10276"  
/note="Isolated from domestic swine in Hartington,  
Nebraska on March 12, 1999"  
complement(284. .736)  
/gene="SPV001"  
complement(284. .736)  
/gene="SPV001"  
/note="Identical to SPV150, similar to lumpy skin disease  
virus LSDV001 and vaccinia virus strain Copenhagen B15R"  
/codon\_start=1  
/product="SPV001 A52R-like family protein"  
/protein\_id="AAL69740.1"  
/db\_xref="GI:18448494"  
/translation="METNIYRNDFLDNDDITALLDLYVHSCISYRHRDPAGKLEFV  
FESPKDAEFVGENIIGFVMEFDSIHGFAHKSMSITMLKKENYIRESCAVVGLL  
ARAAASYGGDERPTASPKVLLKELTDNELFKVKTAAIVRLKKYG"  
complement(811. .1323)  
/gene="SPV002"  
complement(811. .1323)  
/gene="SPV002"  
/note="Identical to SPV149, similar to lumpy skin disease  
virus LSDV002"  
/codon\_start=1  
/product="SPV002 hypothetical protein"  
/protein\_id="AAL69741.1"  
/db\_xref="GI:18448495"  
/translation="MCALSTTRNVFSGINGFVRGIIVYIVKFIGFIYFKSPFQLL  
IDNARIFMLFIFYVDRLLNIITGIYNNIMPMKKFIATLFGNPNPDYIVTKKD  
DDMMTCVYKDKDKDKDNNNNNDIKSSNKTDMNTYTDPAYDFQVLFNFIITPNFRK  
YTEYKCSK"  
complement(1430. .2452)  
/gene="SPV003"  
complement(1430. .2452)  
/gene="SPV003"  
/note="Identical to SPV148"  
/codon\_start=1  
/product="SPV003 MHC class I alpha chain-like protein  
P32231"  
/protein\_id="AAL69742.1"  
/db\_xref="GI:18448496"  
/translation="MITKAIIVLSITAYVDASAFVLYNYVTYVTLQDDNHRDYDFEVTDY  
FNDILIKRLKLNSETGRPELRNPTWNETKIRYYPKNYNFWMFLNMRSETLDEIN  
KLPTSPNYKMSLTIGCTDLRQLQNFQVTVGGINWTRFDPKPKRSRSPFK  
VCMITVKSQHWERWHEHLSGWSVTLTCPTADDYKISKGYIDKPKPVTYVIGIERGD  
NTTLICTFDNHPSSVAVKNYDIEDFADPYRDYVYNELLPTDYLPGEPGPTITRR  
LGDKYLFETSSRPVAVPTMSNIACVGFHSTLEPSIYRCVNCSGPEPVLYQVGGRRND  
LEDEED"  
complement(2494. .2772)  
/gene="SPV004"  
complement(2494. .2772)  
/gene="SPV004"  
/note="Identical to SPV147, similar to lumpy skin disease  
virus LSDV004"  
/codon\_start=1  
/product="SPV004 hypothetical protein"  
/protein\_id="AAL69743.1"  
/db\_xref="GI:18448497"  
/translation="MLSYIINPLLSIVYFILGNVSKLETYILMTIMFIPLRAVPYSL  
ISNGWLSLDSINPFRKRESFLSSLNPRKEETKKKEGFFSGWFG"  
complement(2821. .3630)

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene











```
CDS
/organism="uncultured proteobacterium EBAC31A08"
/db_xref="taxon:133804"
/clone="BAC EBAC31A08"
1593..2807
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted threonine dehydratase"
/protein_id="AAG10439.1"
/db_xref="GI:9971877"
/translation="MPITPEIKDKVRFGAKILQHGDNVDAALKEALFLAKKKKLV
EVHPDDPLTAGOCTIGEILEDKNFDVVPVPGGGGILAGYSAWIAONNKKIKIV
GVEVDSACLAENKANKRVILKEVGLPADGVAVSRVGKNFDFVKECDVSVISVD
EVCVAKDIFEDTRVLSPPAGALAGLAKYARKVKKLLAISGANVQRQLNFIV
ERSEIGENREKLSIKPEIPGSLKSRMGSSQVTEFNYSKLSLSLDAVYLVGVRTK
TEKSPFELKAGFTSDFTRNEISNDHLRMVGGNSDGSNNHNERIFRGEFFE
KPGALLNLEKFGKWXISLFYRNLSGAFGKILIGIESKDKDLINHLNKSQTIFE
ETSNKAYKDLK"
complement(2812..5097)
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted secreted protein"
/protein_id="AAG10440.1"
/db_xref="GI:9971878"
/translation="MKITLASLSLLTSLVLSNSIVIDGSLDEPWNNAPEITDFYE
TSPTFLKNEILLTALIFSNKDIYVGFKNFSQNSDMFSRKTLDQINSLSDKNSINI
DFDGNKFAIFLAINGLNSLDAIKTAQDGFCTDWDGDWIAKTQKFDGYWYSEFFIPW
DVTLMQDPGDIRKINYALRYSAENSWSSAGTMSRNDYFOELDSLEIQNTYKSK
LTFPPYISSNKNVTSNQSDDIGAEIFYSTGKGROIINAIINDFGOAESDDVAINFSA
QDTFSERARFKNQSFDSINVDYRSVINTRRIGAAPSINCSSESENEEGRAAKN
YSDIDFALRQKNSENEFEVASEADEAFSVGRDYSALRSRKLGNKTLGFMVTVN
ENMGESTVNVIDYINIRSKLITYSDLLASERDKNGLGYRTQFYLPNKLNSIS
GSLLYFEEDFLNDEGYLQRDWHAGIGGNRKINEFDSNLSLKQIDYCIDNLADAT
AGNSPIGIDOKNSFSFSDKNSFOLDNFRSSGKWTITTRKNEAVPFIKSRKRIGITA
DEANYSFWVDWVSFEKEDKNRWDNGYGRKFYIAGSIPEPNDLRINAQVIR
KENALYQDNNLASDSRQDTLSFDLWFKNNKHEIRLKSQFVALKAKNPISFSD
KENALYKGNINVEENTGVASPIRYKELAPLSYLFYLSVSGGRIYDEEDRSOSEL
FREPNWPSDELFSIKFLKY"
complement(5126..5701)
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted 5-formyltetrahydrofolate cyclo-ligase"
/protein_id="AAG10441.1"
/db_xref="GI:9971879"
/translation="MVKNKIRKSLFEOQSISDLKSNTSFLIOENVLEKINTHEMQRN
IALLYPFREVNNTLLIOAFTKLKNKXIPKVIDDEMAFNLLDEQPNFFINOFGIKE
VKSEDIYVNIEDLMFIMPTGVDLNGYRLGYSGYFDRLIGSINEVSRPLVLGLCYO
YOICNNHFEKRDHLKIDYVFSNDRMISTSS"
6433..7728
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted xaa-Pro aminopeptidase"
/protein_id="AAG10442.1"
/db_xref="GI:9971880"
/translation="MEKIIFKNRRDSLIKHPKNSALIVPGADLYRNADSYNLRQE
SSFYVLSGFCPSPLLQVNGKSIDSIFFVPEKDKLPEIMDYGVRAGPEKSIDIDIG
QAFENKSDALMPEILQGLEKVFYSIGKKNKGDKQYIDWTCANESKDRHSKSIDIDIG
SNWGNRLRIKDKHEIDMKRACEISAESYIEVMSIKPGDNEQIEALFLYEFKRG
GRPAPYTVAGGEGACVLHYIENDKELASDILVDAGCEYKMTASDITRFFVSGR
FSDQLQIYINHYKANLAAIDAVKTSIMEQMVSEKVIETGLVEGLILSDVNLQHL
KNAGFKDYMHKRWHLGDLVDHVDGVMEGDEFMKFKPMGMITTEIPGIYISSAMDVD
KNGKGIREDDILVTDSGNINLTKVPSNPQIEISLMA"
7732..8892
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted quinone biosynthesis monooxygenase"
/protein_id="AAG10443.1"
/db_xref="GI:9971881"
/translation="MEVPVISGGGIIGNYISIRLEKNNIKTVIVEKASSFKALDKGI
```

```
CDS
/RTVLTNEHSMQMLKNIQPSIAOINSIDVLDRGGTGKIOFLAKDVGSENLSYVTFYN
ELQKLSDPCKERTLFNNEIDSVQNLNATESDPEIMLDGMTIKTNLAGCDGRNSNIA
KIASLTSSFDYLTALTFTVVDINDSDSHGRAHQVFSKGIAPALMPLEGKGMNKCIV
VMSIKNOVLGDEPVSEFNKNISSEFKLNVSRLKSEILSFKLSNHHFENYISGPV
LLGDAHSIHPLAGQINLGADATFCEEVSIYKKGIAFNEKSVLRKYRVEIRKSMN
FLMLKSMDFVDLFGSENLYLRLIRNLGISLNSKSKFVKAFFFIRHASMKNKF"
complement(8900..9136)
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted 50S ribosomal protein L28"
/protein_id="AAG10444.1"
/db_xref="GI:9971882"
/translation="MSKIQVTKIPOSNNVSHAKNRTKRKFFPNLHTHKFWVESEN
RFVNLKVSAGKMRIDIRKGISSEVLKELRAKQKV"
9710..10129
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted deoxyuridine 5'-triphosphate
nucleotidohydrolase"
/protein_id="AAG10445.1"
/db_xref="GI:9971883"
/translation="MGSKIPLPQVETKGSAGLDRACLDSNLSLQAGTSOLIPGFAM
YLEDPLGAAMVIPSGLSGKHGIVLGNLVGLDSDYOGELVPAWNRSDTDTEINPGD
RIAQMIVPVIOADFVIEVDEFNETQKGEKFGSSGIN"
10443..12431
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted primosomal protein N' (replication
factor Y) - superfamily II helicase"
/protein_id="AAG10446.1"
/db_xref="GI:9971884"
/translation="MKIFYDIAYSLPLRQCFTYSSELKITKTRVSVPPFGKRKIVG
VIKNIQPDFLKAKAIKKIITAVLDEYPLDKPIFDSILWSSDYVHHPIGVNFTFIP
TELKINNNKTEALREFSEYSVNEDDKFDLTQDEKAVKALSXKSGSPPLLXGVTC
SGKTEYLRVAETFIKNKNSVLVLPEINLTPQLLSRFENFNGEIGIYHSHKQTAAR
LKTWLKAFSGIKIIVGTRSSALVPLDNIGLIIIDEEHDQSFQSGEFGFKSARDLSIK
RAQLADIPILIGSATPSLOTILKVENKFIIRDVFNVDGNKPKLIALDINNPSPLIG
VAVKETIEAMQSTIDRGEQVLVFNRRGFAPLYQSCSGWADCKSCDTNLFVHOARN
RLICHRCEASVSNLSCPACKNDFNMYAGTERVEEVKSSVFKVPTPIRVDHDSHTK
VGAMEAIVKKIHSDDAAILVCTOMLAKGHDPKVTLSVLNADNGLISPEINALEKIS
QLLQVSRAGRNNNNLKVIIQTRYPDIDNINLKTGYMKFASCSLSTNQMNLPPIF
TTLCLLSCSPTQKSNVDFLEKAVLILNSRTDINVGIPPLSLVSKSGNYRHQVITIHA
PKTFLNKLKFLTEFEKWPESKNKWKWSFDIDPIDLS"
complement(12428..13510)
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted N-acetylmuramoyl-L-alanine amidase"
/protein_id="AAG10447.1"
/db_xref="GI:9971885"
/translation="MEINOSSLPTEINVPYPIKKVRASODGSLTRIVVDLYESVHW
ONPQTINTENIKLELVKRNKLNKSIRDIVVADAGHGKYPGAVGPNINLEKDVIT
LLIAELBERTLDIYGYRPMIRGDDETLIDNNRQDARKYGADIFVSIHADGPRLS
VKGASVTSWSEADSVARNLSKQORERIQADIKNLKDPDFDEDAARTYPEIYKKKI
SESKTGLKFIILDLQRDPFTKIHKKNVETADFRVLKSIDIPSVLSESGFITNPDAQR
LKGKPRGMIAARSVFLGTHNYFKDKPKANTFMSIDPGFVTKYIQKGDVLSLTAIRFGV
TVEETENKLNKRPYPGOLIKINI"
complement(13670..14131)
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="predicted kinase of the phosphomethylpyrimidine
kinase (Thid) family"
/protein_id="AAG10449.1"
/db_xref="GI:9971887"
/translation="MKKLTILNDEATNQLSGKIAMIEILKSSQIEIHLLEGDLGAGKT
FISRIIKNCGMKDLVKSPYTLTCEEYDFNNLMFLIDLYRTNEAEDIDIFLSRKIN
SKVYLIIEWPERLOHRSFDLKIIIFSHLPEGFEVSLISGDNFKEWLKYE"
complement(14128..14985)
/codon_start=1
```







```
seq_name: gb_pat:190034
seq_documentation_block:
LOCUS I90034 918 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 3 from patent US 5723436.
ACCESSION I90034
VERSION I90034.1 GI:3409974
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 918)
AUTHORS Huang,L. and Cyert,M.S.
TITLE Calcineurin interacting protein compositions and methods
JOURNAL Patent: US 5723436-A 3 03-MAR-1998;
FEATURES
Location/Qualifiers
source
1..918
/organism="unknown"
BASE COUNT 313 a 201 c 163 g 241 t
ORIGIN

alignment_scores:
Quality: 74.50 Length: 95
Ratio: 1.330 Gaps: 2
Percent Similarity: 58.947 Percent Identity: 26.316

alignment_block:
US-09-528-682-4 x I90033 ..
Align seg 1/1 to: I90033 from: 1 to: 1222
13 ProlleAlaAsnThrAspThrThrGlnPheLeuAsnMetGlyAsnG1 29
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 CCGAGTGAGACGTTAATTCACACACTCTCTATTTTGCAGCTTTAAAC... 495
29 uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluG 46
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
496 .ATGAGTCTTGATGGAGAGAAAAAATCGTCCAGTCTACACGCAACAT 544
46 lyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
::: ::: :::
545 CAGTGTACCTAGTACAATAAGACAGACAAT..... 576
63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
577 .....CAGCATTTCAATGACATATAACAGATGCTAGGAGTAGTGACGA 620
79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
621 AGATGCCTTTCCAAAAGCCAATCATTAATTTCAATAAGAAACTACCAA 670
96 euLeuIleAsnAsnAspGlyPhePheSerAsnAsn 107
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
671 TACTTAAATAATGATGATAACGTCATACATAATCAAAAC 705

seq_name: gb_pat:190036
seq_documentation_block:
LOCUS I90036 2745 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 6 from patent US 5723436.
ACCESSION I90036
VERSION I90036.1 GI:3409976
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 2745)
AUTHORS Huang,L. and Cyert,M.S.
TITLE Calcineurin interacting protein compositions and methods
JOURNAL Patent: US 5723436-A 6 03-MAR-1998;
FEATURES
Location/Qualifiers
source
1..2745
/organism="unknown"
BASE COUNT 915 a 587 c 495 g 748 t
ORIGIN

alignment_scores:
Quality: 74.50 Length: 95
Ratio: 1.330 Gaps: 2
Percent Similarity: 58.947 Percent Identity: 26.316

alignment_block:
US-09-528-682-4 x I90036 ..
Align seg 1/1 to: I90036 from: 1 to: 2745
13 ProlleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnG1 29
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
577 .....CAGCATTTCAATGACATATAACAGATGCTAGGAGTAGTGACGA 620
79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
621 AGATGCCTTTCCAAAAGCCAATCATTAATTTCAATAAGAAACTACCAA 670
96 euLeuIleAsnAsnAspGlyPhePheSerAsnAsn 107
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
671 TACTTAAATAATGATGATAACGTCATACATAATCAAAAC 705

seq_name: gb_pat:190033
seq_documentation_block:
LOCUS I90033 1222 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 1 from patent US 5723436.
ACCESSION I90033
VERSION I90033.1 GI:3409973
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 1222)
AUTHORS Huang,L. and Cyert,M.S.
TITLE Calcineurin interacting protein compositions and methods
JOURNAL Patent: US 5723436-A 1 03-MAR-1998;
FEATURES
Location/Qualifiers
source
1..1222
/organism="unknown"
BASE COUNT 313 a 201 c 163 g 241 t
ORIGIN
```







1302 GGGTAATGATCTGAGGCCAAAAATGTATCGTAGTGAAATGGTAAATGGTG 1351

101 spGly.....phepheSerAsnAsnGlyGlyLys 110

1352 TAGGTGTTGTTATTTTATGTTTGGAAATAGAGCAAGGAAA 1392

seq\_name: qb\_pl:SCYKR021W

seq\_documentation\_block:  
LOCUS SCYKR021W 3338 bp DNA linear PLN 11-AUG-1997  
DEFINITION S.cerevisiae chromosome XI reading frame ORF YKR021w.

ACCESSION	228246	Y13137
VERSION	728246	1 GT:486442

VERSION Z28246.1 GI:486442

## KEYWORDS

SOURCE baker's yeast.

ORGANISM  
Saccharomyces cerevisiae  
Saccharomycetes cerevisiae  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

## REFERENCE

**AUTHORS**  
Duesterhoeft, A., Moestl, D., Poehlmann, R. and Philippsen, P.

**JOURNAL** Unpublished

REFERENCE 2 (bases 1 to 3338)

**AUTHORS**      **MIPS.**

**TITLE** Direct Submission

**JOURNAL** Submitted (09-MAY-1994) Data collected by MIPS on behalf of the European yeast chromosome XI sequencing project. MIPS at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152 Martinsried, FRG; E-mail: Mewes@chemie.mips.biochem.mpg.de

## FEATURES

source

/organism="Saccharomyces cerevisiae"

/db\_xref="taxon:4932"

/chromosome="XI"

CDS 402. . 3149

```

/notes="QBE

```

```
/codon start=1
```

```
/contig_id="C
/protein_id="C
```

```
/protein_id= CAA82093.1
/db xref="GI:1866443"
```

Odd-SSlms"jex qp/  
qQ\_Ix\_qr/  
GI:480443

```

/uu_xref= SWISS-PROT:P3011/
/translation= MIENTENDTVADV

```

[illegible]

BASE COUNT	1097 a	715 c	582 d	944 t
------------	--------	-------	-------	-------

## ORIGIN

alignment scores:

Quality: 74.50

Quality: 74.50  
Ratio: 1 330

OTJPN 58 047  
055 T

PERCENT SIMILARITY: 38.34 / PERCENT IDENTITY: 20.310

alignment block:

US-09-528-682-A y SCVK8021W

Align seq 1/1 to: SCYKR021W from: 1 to: 3338

13 proIleAlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGln 29

[illegible]

2676 CCGAGTGAGAACGTTAATTTCACACACTCCCTATTTTGCAGTCTTTAAAC... 2723



```

96 euLeuileAsnAsnAspGlyPhepSerAsnAsn 107
:||||| |||||::: :::::|||||
2873 TACTTAAATAATGATAACGTCATACATCAAC 2907

seq_name: gb_in:L14324

seq_documentation_block:
L14324 38269 bp DNA linear INV 09-AUG-2001
Caenorhabditis elegans cosmid ZK112, complete sequence.
L14324
L14324.1 GI:289740
HTG.
SOURCE
Caenorhabditis elegans.
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 38269)
The C. elegans Sequencing Consortium.
Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science. 282 (5396), 2012-2018 (1998)
99069613
REFERENCE
2 (bases 1 to 38269)
Du, Z.
The sequence of C. elegans cosmid ZK112
Unpublished
3 (bases 1 to 38269)
Waterston, R.
Direct Submission
Unpublished
4 (bases 1 to 38269)
Waterston, R.
Direct Submission
Submitted (09-AUG-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: rwenematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate chemistry
or covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by sequence from
more than one m13 subclone.

NEIGHBORING COSMID INFORMATION

The clone sequenced to the left is C03B8, 200 bp overlap; the clone
sequenced to the right is ZC97, 1200 bp overlap. Actual start of
this clone is at base position 1 of CELZK112.

NOTES:
Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in preparation).
Location/Qualifiers
1..38269
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="III"
/clone="ZK112"

FEATURES
source

```







This sequence is the entire insert of clone RP11-78J21. The true left end of clone RP11-93H24 is at 157212 in this sequence. The true right end of clone RP11-245D16 is at 50244 in this sequence.

FEATURES	Location/Qualifiers	Source
	1..163390	/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="13"
	/map="g32.2-33.3"	
	/clone="RP11-78J21"	
	/clone_lib="RPC1-11.1"	
	360..681	/note="AluY repeat: matches 1..311 of consensus"
	682..729	/note="24 copies 2 mer at 100% conserved"
repeat_region	1368..1676	/note="AluSg repeat: matches 1..309 of consensus"
repeat_region	1697..2811	/note="L1PA10 repeat: matches 5052..6164 of consensus"
repeat_region	2819..2898	/note="L1P3 repeat: matches 1065..1144 of consensus"
repeat_region	3522..3929	/note="L1ME3A repeat: matches 5736..6159 of consensus"
repeat_region	4000..4297	/note="AluSg repeat: matches 1..298 of consensus"
repeat_region	4315..4521	/note="L1M4 repeat: matches 4798..5007 of consensus"
repeat_region	4522..4829	/note="AluJo repeat: matches 1..298 of consensus"
repeat_region	4830..5180	/note="L1M4 repeat: matches 5007..5368 of consensus"
repeat_region	5181..5459	/note="AluSg repeat: matches 1..301 of consensus"
repeat_region	5460..5670	/note="L1M4 repeat: matches 5368..5581 of consensus"
repeat_region	6159..6630	/note="MER77 repeat: matches 61..565 of consensus"
repeat_region	6860..7064	/note="L1M4 repeat: matches 4266..4470 of consensus"
repeat_region	7131..7326	/note="L1M4 repeat: matches 4689..4894 of consensus"
repeat_region	7413..7480	/note="MER21B repeat: matches 725..792 of consensus"
repeat_region	7773..8174	/note="CpG island"
misc_feature		/evidence="not_experimental"
repeat_region	8435..8715	/note="AluSx repeat: matches 6..286 of consensus"
repeat_region	9060..9264	/note="AluJo repeat: matches 1..206 of consensus"
repeat_region	9265..9374	/note="L1MB6 repeat: matches 5773..5909 of consensus"
repeat_region	9385..10324	/note="L1M4A repeat: matches 5319..6299 of consensus"
repeat_region	10335..10493	/note="AluJb repeat: matches 131..289 of consensus"
repeat_region	10510..10624	/note="AluJb repeat: matches 184..312 of consensus"
repeat_region	10626..10894	/note="AluJo repeat: matches 1..298 of consensus"
repeat_region	10895..11922	/note="L1M4A repeat: matches 4289..5331 of consensus"
misc_feature	11746..11812	/note="Sequence from overlapping clone RP11-245D16(ALJ59513). Assembly confirmed by restriction digest."
repeat_region	11923..12216	/note="AluSx repeat: matches 1..295 of consensus"
repeat_region	12217..13237	/note="L1M4A repeat: matches 3254..4289 of consensus"
repeat_region	13238..13542	/note="AluJo repeat: matches 3..304 of consensus"
repeat_region	13543..13928	
		/note="L1M4A repeat: matches 2864..3254 of consensus"
	13929..14219	/note="AluJo repeat: matches 1..288 of consensus"
repeat_region	14220..14776	/note="L1M4A repeat: matches 2397..2864 of consensus"
repeat_region	14777..15079	/note="AluSg repeat: matches 1..302 of consensus"
repeat_region	15080..15600	/note="L1M4A repeat: matches 1869..2397 of consensus"
repeat_region	15601..16623	/note="L1M4 repeat: matches 4015..5093 of consensus"
repeat_region	16618..17254	/note="L1M4A repeat: matches 251..882 of consensus"
repeat_region	17257..17563	/note="L1ME repeat: matches 5281..5602 of consensus"
repeat_region	17580..17741	/note="L1M4 repeat: matches 6129..6294 of consensus"
repeat_region	19008..19077	/note="35 copies 2 mer aa 65% conserved"
misc_feature		complement(19367..19787)
		/note="match: STS: Em:G49126"
repeat_region	21114..21838	/note="MER11A repeat: matches 349..1070 of consensus"
repeat_region	24587..24666	/note="40 copies 2 mer ta 63% conserved"
repeat_region	26010..26349	/note="170 copies 2 mer at 62% conserved"
repeat_region	26094..26221	/note="4 copies 32 mer 79% conserved"
repeat_region	26109..26348	/note="5 copies 48 mer 65% conserved"
repeat_region	26380..26559	/note="TIGER1 repeat: matches 1..181 of consensus"
repeat_region	26560..26845	/note="AluSg repeat: matches 1..287 of consensus"
repeat_region	26846..27856	/note="TIGER1 repeat: matches 181..1182 of consensus"
repeat_region	27860..27991	/note="L1M3e repeat: matches -667..-532 of consensus"
repeat_region	28699..29605	/note="L1M3e repeat: matches -487..-439 of consensus"
repeat_region	29679..31092	/note="L1M2 repeat: matches 705..2572 of consensus"
repeat_region	31019..31117	/note="L1 repeat: matches 2118..2215 of consensus"
repeat_region	31118..31417	/note="AluY repeat: matches 3..302 of consensus"
repeat_region	31418..32834	/note="L1 repeat: matches 2215..3639 of consensus"
repeat_region	32840..33350	/note="TIGER1 repeat: matches 1911..2417 of consensus"
repeat_region	36375..36672	/note="AluSc repeat: matches 1..305 of consensus"
repeat_region	37931..38257	/note="L1PA1A repeat: matches 5819..6149 of consensus"
repeat_region	38258..38558	/note="AluY repeat: matches 1..299 of consensus"
repeat_region	38559..39059	/note="L1PA1A repeat: matches 5332..5819 of consensus"
repeat_region	40071..40341	/note="AluSx repeat: matches 1..271 of consensus"
misc_feature		complement(40499..40598)
		/note="match: STS: Em:G43515"
repeat_region	41721..41977	/note="AluSg repeat: matches 33..289 of consensus"
repeat_region	42003..44320	/note="L1M1 repeat: matches -152..2217 of consensus"
repeat_region	44321..44628	/note="AluSc repeat: matches 1..308 of consensus"
repeat_region	44629..44901	/note="L1M1 repeat: matches 2217..2484 of consensus"
repeat_region	44906..45078	/note="L1M2 repeat: matches 3524..3690 of consensus"







```

AMDYTVVVAAGSEPATLQFLAPYTGATIAEYFMYGKATLVIVYDLSKQAOYRQMS
LLRRPPREAYPGDVFIHSRLLEAAKLSDELGKSGMTALPIETIQAQDVSAV IPT
NVIISTDQOIFLSSDLNAGIRPAVNCISVSRVGSAAQTAMKKVAGIKLEAQPD
DLQAPAFASDLKATQDLARGORLRELLKQSONQPLSVAEQVAAILFAGINYYDDI
PVDKVTFTKGLRDYLSGVNPNFYFQDVQSKALGDDEEKALKAALEDYKTKTKATA"
complement(6138. .6689)
/gene="atpb"
complement(6138. .6689)
/gene="atpb"
/note="ORF_ID:all0006
atpb gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit delta"
/protein_id="BAB77530.1"
/db_xref="GI:17134984"
/translation="MTSKVANTEVAQPYAQAALLSIKSKSTEEFGTDARFLNLLTE
NQQRNFIDNPFIAENKALIKQILSEASPYLRNLLLVLDKRRIFLPLPILQOYLA
LLRLQNTVLAEVTSVALETDDQQAUTEKVLTAKRQVELATKVDSDLLGGVLIKV
GSQVIDSSIRGOLRRLSRLSNS"
complement(6686. .7249)
/gene="atpb"
complement(6686. .7249)
/gene="atpb"
/note="ORF_ID:all0007
atpF gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit b"
/protein_id="BAB77531.1"
/db_xref="GI:17134985"
/translation="MGTFLLMAEASAVGELAEGBARGFGLNTNILDNLINLAI
ITLVFGRKVLGNTLTKRRENIETAKNAQRADAKQLKEAQKLEQAQAEERI
KKSADNAQTAGAILIQAADVIERLQAGAADLNAELDRAIQLRQVVALQKVE
SELGGISDAQKTLIDRSIAQLGGV"
complement(7359. .7850)
/gene="atpc"
complement(7359. .7850)
/gene="atpc"
/note="ORF_ID:all0008
atpG gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit b"
/protein_id="BAB77532.1"
/db_xref="GI:17134986"
/translation="MTHWITLAVEKVAKEGLFDLDTPLMAIQFLLALILNATL
YKPLGKAIDGRNEVYRNQLEAQLRSKAERLAELAEYEQELAGARRQATIIADAQAEA
QKIAAEKVAQAQEAQAEQAAGEIEQKQALASLEQVQVDALSRLERLKLGLADLV
KQR"
complement(7996. .8241)
/gene="atph"
complement(7996. .8241)
/gene="atph"
/note="ORF_ID:as10009
atph gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit c"
/protein_id="BAB77533.1"
/db_xref="GI:17134987"
/translation="MDPLVSAASVLAALAVGLAAIGPGIGGNAAGQAVEGIARQPE
AEGKIRGTLLLSLAFMEALTYGLVVLVLLFANPFA"
complement(8324. .9079)
/gene="atpi"
complement(8324. .9079)
/gene="atpi"
/note="ORF_ID:all0010
atpi gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit a"
/protein_id="BAB77534.1"

```

```

/db_xref="GI:17134988"
/translation="MLNLFNFVSVPFLAEVYKHLWIOGNLKLHGOVFLRSWFVIGV
LYLVAASSNVKRPISQIONLLEVALEFIRDLAKNQIGEKERYEPWPVFGTFLFIF
VSNSGALVFPKLHLPGEGTAPTSDINTVVALLSLAYFYAGFSKGLGYPGNY
VQPVSAIIPFKIIEDFTKPLSFRFLGNILADELVGVLLVLLVFLFVPLPVMALGLF
TSAIQALFATLAAAYIGEAMEDHHGESEHH"
complement(9132. .9500)
/gene="atpi"
complement(9132. .9500)
/gene="atpi"
/note="ORF_ID:all0011
atpi gene product"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit 1"
/protein_id="BAB77535.1"
/db_xref="GI:17134989"
/translation="MQEYFYQELVLTILVLTVGVFVSVWIFYSLNIALNYLGA
GVVYLRMLAKDVERLGRKQSLKTRLLMALILLASRWNLQIMPIFLGLFTYKAT
LIIVVRVAFISDSPLKRP"
complement(10130. .11329)
/gene="all0012"
complement(10130. .11329)
/gene="all0012"
/note="ORF_ID:all0012
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB77536.1"
/db_xref="GI:17134990"

```

```

alignment_scores:
    Quality: 74.50      Length: 129
    Ratio: 1.112      Gaps: 6
    Percent Similarity: 51.938      Percent Identity: 23.256

alignment_block:
US-09-528-682-4 x AP003581 ..

Align seg 1/1 to: AP003581 from: 1 to: 348050

12 GluProIleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAs 28
|||||:|||||:|||||:|||||:|||||:|||||:
303585 GAACCTTAAGTGTACCAACACC.....TTACGGGGTGTGA 303622

28 nGluValAla.....LeuAspGlyArgTySerAsnTyrA 40
: ||||| ||||| ||||| |||||:|||||:
303623 TACCGTTGTGTAGTATCACCGCGCTTTTAGACCAACGCTTTGAAGGATACC 303672

40 laLeuIleSerAlaGluGly..... 46
|||:|||||:|||||:|||||:|||||:
303673 GCGTTCNAACCTCACTGGCGGTAGACTTTACACGACCAATCCTCGTCCA 303722

47 GlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProG1 63
|||:|||||:|||||:|||||:|||||:
303723 GCAACCAACACAGATGAGCGGTACACTCAAGGTTGCCAGCTTTAACGT 303772

63 uValArgGluPheAsn.....SerLeuProA 72
|||||:|||||:|||||:|||||:
303773 TCTGAATTACTTCAACGGTGATGCTAGTGTAGCGGCTTCCACGAGTCTGTG 303822

72 snAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAsp... 87
|||||:|||||:|||||:|||||:|||||:
303823 ACAAGGGGTGCAGAAACCTGACTGAGTTCAACCCCGCGTGAAG 303872

88 .....AlaAspPheLysLysTyrIleLy 95
|||||:|||||:|||||:|||||:
303873 ACAATTCGGGTATTCTGGGGTTAAATGCCGATGTTGTAGGTCTGATTGA 303922

95 sLeuLeuIleAsnAsnAspGlyPhePheSerAsnAsn 107
: |||||:|||||:|||||:|||||:
303923 A.....ATCGAAATGATGTTACGGTGTGTAATAGT 303953

```











```

CDS
/gene="Atul683"
122..346
/note="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL42683.1"
/db_xref="GI:17740118"
/transl_table="MLKHGALIKNORQMFVREKQSTSAKISVISTAWTVLVIAWVA
TFFSLYDKPKPQAGPDYQLTARYQYMY"
complement(373..555)
/gene="Atul684"
complement(373..555)
/gene="Atul684"
/note="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL42684.1"
/db_xref="GI:17740119"
/transl_table="MLKSAEADPPRESASVSTRLEFLILTPHLGGDDGHRHEHYA
DQQRKORCEHDIVS"
625..777
/gene="Atul685"
625..777
/gene="Atul685"
/note="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL42685.1"
/db_xref="GI:17740120"
/transl_table="MTAEKPPSNHTSVKTPFFQKFPFSISLWIDLRNADTLRAGLA
TAMTFL"
791..2683
/gene="ppid"
/note="Atul686"
791..2683
/gene="ppid"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl cis-trans isomerase D"
/protein_id="AAL42686.1"
/db_xref="GI:17740121"
/transl_table="MDSLRNARSMTAAKLLLLLVVFGVWGSASLVTANSHAVMT
VGDQTVSQEFLAYQORVSDLSRQFGLTTEQAKAFGIDRQVFSQAAGAALDELA
SKMNLGLSENRLAIYEDDPAFKSVNGQFDRNLFSERLNSGFREDDYIKERSKVAVR
SOIVASVDGFAAPQVLDALKQYRNEQRAVDYVILSNVIPPVKAPGDDVLTWPFTF
NKSXRAPEPFKFTVVKLEPSSDIAEPASVTDQAIDYVNSHKDSFPTAGRTVEQLTF
PKEMAAAAQIRLQNTYDQVVKDQKTSADSVLTGFTKDTIPDQSIADAAFAIQK
DGVSPVVDGSGPILLRVTKIPETTLTDEAKEDIRKDLATAAAAEIINVHDIYE
DLRAGSLADAKNLNKPVTIDAIDAGLNEKGDIAEGLPSPOLAQDVFKTEPGTD
ALPILNGREGYIWFQVEQIIPARDRTLAERVDVADWTAEQORTALAAKADLKARV
EKGETLEAVAGELSIAVQKSGRLRTSEDAIFGROTIAAVSGAGVVGTAADAGSS
RILFKVTVSDANAPADALANDDDQFTARAAGDDMLDMVNRQLNDIGVTINQALAD
QAWVG"
2698..3726
/gene="trpd"
/note="Atul687"
2698..3726
/gene="trpd"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="anthranilate phosphoribosyltransferase"
/protein_id="AAL42687.1"
/db_xref="GI:17740122"
/transl_table="MAELKPLIAIKVANGESLNREDARTAFDILMSGEATPPOIGGFLM
ALRVRGETVDEIVGAVSSMRMLPVSAPANAIDIVTGGDGIQTYNISTASIITAG
TGLPVAKHGNRALSSKSGTADALSALGVRLDIGPDLIARCIAEAGLGFMEFQAQMHSAM
RHVGPSRVELGTRTIFNLPLSPNAGAKRQLGLVFPWLVLPVLAELVRLDGSGSIWV
VHGDGMEVTTTGVTHVAALGDKIRTFDLTPKDFGVEPALMNDLKGGDGTANAALR
EVLGSKRNAYRDISLCNNAALVIAGKAEITLSQAMTIASDALDSGKAAALDRLVAYS
NEANSQQE"
3737..4552
/gene="trpc"
/note="Atul688"
3737..4552
/gene="trpc"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="indole-3-glycerol-phosphate synthase"
/protein_id="AAL42688.1"
/db_xref="GI:17740123"
/transl_table="MSDILKRIETYKLEETAAAKAVKSLADLKAMADQSDAPRGFYKA
LRKQAAGKFGLIAEIKKASPSGLIRPDPDPALAAAYEAGAACLSVLTPTPSFG
APEFLTAARNACALPALRKDFMFDYQVHEARAWGADCILIMASLSDDEAKRLEDEA
FALGMDVLIIEVHDAEETERALKLTSPLLGINNRNLTFEVGLTSEKLAGLVPADKLL
VGESGIFTEFCQRLKESGISTFLVBESLMRKDDVTAATKALLTGHSGILAAE"
4560..5051
/gene="moac"
/note="Atul689"
4560..5051
/gene="moac"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="molybdenum cofactor biosynthesis protein C"
/protein_id="AAL42689.1"
/db_xref="GI:17740124"
/transl_table="MSDQKLTHTDASGAHVVDGDKAETRVVAVAEFGVKMKPETL
ALIRDNAKGDVIGTARLAKIGAMAKOTANLIPCHLMLTKVAVDITDGLPGLRV
EAMVKLSGKTGMEALTAVTAVSIACLTIDYMAKADKGMIEIVNIRLLEKSGKSGDFRR
QES"
5052..6269
/gene="moa"
/note="Atul690"
5052..6269
/gene="moa"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="molybdopterin biosynthesis protein"
/protein_id="AAL42690.1"
/db_xref="GI:17740125"
/transl_table="MNPLLPVDEAIKRLLEAAVPVTDSETLPLAECGRVLATLAAR
LTOPPDSAMDGVALBSADAAEIGSVLTILGQAAAGHAFAGTVGKEAVRIFTGAPV
PDGADTILIQDEALDDGKTRTFVAVTAGRHRIIPRGQDTEGETALEAGRELGTDL
TLAAMNHAALSVYRRPRIALATGDELLPPGSAPEAQIIIASNTFVGAALARAGAE
IADGIIADDDSLIAAARRAVAGVDVLTGGASVDGHDLDLQVAAKAEQMQLDFWR
VIMRPGPLMGVAGTGVQLGPNPVASLVCGLIFLEPLIRRIARRAPLQRSTART
ATPLKANDORODYLRARESADENGALVAEYAKAKOSSMMKTLAHSDDGLIVRPPHAE
APGATCPVIRLKP"
6751..7821
/gene="Atul691"
6751..7821
/gene="Atul691"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="GDEF family protein"
/protein_id="AAL42691.1"
/db_xref="GI:17740126"
/transl_table="MNTASTPTKTPDIAGOIAYAMRSMGVSPIPNYSFLVEAYIGS
NPALTAKLAALGNBVTQDDLSLSRYGEGNPNBSIDDAOKDLRLLEALGTLRKEQ
SSIDENYRILGETQRIDDKNAASNILRNALISLADATGSKIIDGETFTFDVNRHAE
EMHQVRLELDEYKRIANTISLRNRRRAFDRLASVYDSSIGQYITTLVLDDIHFK

```











the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RCP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from SP6 to T7 of the PAC clone. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomesSeq.html>.

## FEATURES

Location/Qualifiers

1..144322

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db\_xref="taxon:4530"

/chromosome="1"

/clones="P0463F06"

join(3116..3947,5188..6259)

/gene="P0463F06.1"

join(3116..3947,5188..6259)

/note="Probably inactive due to frameshift in CDS"

/gene="P0463F06.1"

pseudogene, receptor serine/threonine kinase"

/codon\_start=1

/pseudo

complement(join(8047..9173,9246..9430,9518..10293))

/gene="P0463F06.2"

complement(join(8047..9173,9246..9430,9518..10293))

/gene="P0463F06.2"

/codon\_start=1

/product="putative receptor kinase"

/protein\_id="BAB17115.1"

/db\_xref="GI:11034591"

/translation="MAVAGAFRRSTALQALTALSLAALLQVEGMYGPEAAQTA  
QCPFFSHLKNVSPFRHGDPPDCYKSHELVCSDTKATIRINNATYVYTRIHS  
LVVVDDLLNSHNSCLPRROHOPHERISMAPTSDGVOVIELAFTNRYRIVHS  
SKEITRNVKPVACLSTHSFYVLGSESETVYVLESFCSGLMTPLDDRGASV  
NASUSYADLVKMRKGAIGISSSGTYHDFRACLAESVSIFFQTVVPSDVIKNN  
LLAIIAFLPFWQCIMAPKLVAYEFAIYGTGPAMSIITWIAVLRCFLAPLVILT  
FLSHKYKARITIDAVEKFLRMQMLSPMRYGTDITAITSHFRDLRGQSGYVYVG  
VLLPGNTHVAKMLGNSNGCEEFISEVSTIGRIHVNVVRLVFGFSEEMRRALVE  
YMPGSLDKYIFESDKSFRDKLNEIALGTARGINYLHQCDLAILHFDIKPHNILL  
KNFVPKVDGRLARLPKDSKVPVSAARGTVGTIAEMISRSFGVSSKSDVYSFGM  
LLEMGACRRNADNAENSQAQYPSRVYRQLTQETGEITAAADMELEKLCIVGL  
WCIQMRCDRPMSEVIEMLGGVDCIQIPRPFCDDDYIPAMESLYLSSELELAAI  
SEKDEESITELN"

join(11349..11391,11477..12273,12952..14085)

/gene="P0463F06.3"

join(11349..11391,11477..12273,12952..14085)

/gene="P0463F06.3"

/codon\_start=1

/product="putative receptor kinase"

/protein\_id="BAB17116.1"

/db\_xref="GI:11034592"

/translation="MRTPIPLQLKQIIVLLLSLLNEPYSTKAWDEGFLKSSOR  
CSKHGPIRPFRLSTHDPSCGMQLSCSGHDTILDHPVLGCKVTAIYRYRVIN  
AILLVESSCNCPQLKASTNVSTAVYEPQKDGATVGCCKSDSTATQDQVIGPRSL  
NLSHGSQALVLPPTYMSALPPDCTVAAQIPYNYDRNGNENQYLDLSNLEK  
ANKVINGETATVHLANKITNACQCEWNGHCCGFSORGQAFCQHGOLILHJFEFK  
GILQILAHITLLMEQTNFYFYFTATSSATLVLLGMVGTIYLSLTKRYNEIHLK  
VEMFLYGTSPKPRYSFSEVKLITRRFRKIGHLLNVYKGLPGLVGVAVAKLEN  
SRGQGVFINVEATIGRIHANVLRLLGFCSEGRTRTLIYFEMPNDLEKRYIFSQGN  
VSREFLVPMKMLDISIAGRMGYLHQGCNQRIHLFDIKQNILDYSFNPKIFDGL  
AKLCARQSVITVTAARGTMGYIAPELYSRNFVYSYKSDVYSFGMLVLEMVSGRRNL  
DPGIDNOHEVYFLEWYKVTGONLLIGMETODEKYKVRKLAIVALWCQWPKNR  
PSTTQVVMNLTGRQLDQMPKPFVSSASSIP"

complement(join(14626..14732,14821..15151))

/gene="P0463F06.4"

complement(join(14626..14732,14821..15151))

/gene="P0463F06.4"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB17117.1"

/db\_xref="GI:11034593"

/translation="MCOIGDGRSILFWEDCWMQEGGIRAIAPSIPEHVPVIRKRTV  
AEALLNRAWIRDISGALGQVQAIREYKLKLSLQONAGCNPSQDVICWKMSSGOYSK  
SAYRALFDAGORQIDVYVAVYHILDTAFYVINTTRVLITFLA"

16173..17087

/gene="P0463F06.5"

16173..17087

/gene="P0463F06.5"

/codon\_start=1

/product="putative receptor kinase"

/protein\_id="BAB17118.1"

/db\_xref="GI:11034594"

/translation="MATALYLSLSRYDEEVLKVMELFRTYGTSKPTRYSFSDVKKI  
TRCFEQGGGSGFVSNIVRLIGFCSEGRTHILIYEFNPESLEKRYIFFHDPNTSQEL  
LAPKMLDIALGIARGMYLHQGCNQRIHLFDIKPHNILLDYNFLKISDFGLAKLCA  
RQSVLTUTKARGTMGYIAPELYSRNFGEISYKSDVYSFGMLVLEMVSGRRNSDPSIE  
DQNEVYFLEWYKVTGQDFVLSGEMTEEDRLKVRQMALVALWCIONPRNRPSTWK  
VVMNLTGRQNIQVPPKPFVYESHAVB"

complement(join(17879..18106,18186..18372,18654..18962,  
19113..19291,19329..19461,19516..19624,20367..20631))

/gene="P0463F06.6"

complement(join(17879..18106,18186..18372,18654..18962,  
19113..19291,19329..19461,19516..19624,20367..20631))

/gene="P0463F06.6"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB17119.1"

/db\_xref="GI:11034595"

/translation="MGGOEQRSSCSRSSPSSSSCHSTSSMSATGCVAVGEYBAGQ  
LYDREQVTELARLHCRNPIPLSPPELVYVYVSTPSYVGGELAPGEDVGGLLPSAGQ  
GGRGRRRRPVGTGTMAAGDCAEAYDEIEGRELVGQRPAPGAGHAGKCRAGY  
EAEDEEQASRRRPLGLKLITLAMDVODTLDLDMQMGSGYGGGWTFRDIFFNLHQ  
NQNINQINQSMNTLSSVRLRDVMASSNRSTLAPSAAAIYARLRLPAAADPNAA  
VVNQAVALSAADLRAIGVSARKAAYLHDLAGRAAGELSESAAVATDEAPAPSRQ  
PTLSPAAVPEKSEPERERERKGRVMTWPMWMPGRGSHADSVAVTDKTMKMTTEG  
PRLNLACTFSIPPPPLATNPSPSPAIPTPGHLCLFCGDVSALRSAYNARCPGYASP  
SPCVLRPEPQIQAEPOD"

join(22482..22821,23165..23209,24203..24528)

/gene="P0463F06.7"

join(22482..22821,23165..23209,24203..24528)

/gene="P0463F06.7"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB17120.1"

/db\_xref="GI:11034596"

/translation="MPARCSHRPLRSSAGRADHMLSSMPDPVTTWATDTTAYIMDE  
QOEIKFKTSKWNPIRPARLLTSNGROIYIRPPFSAREYLMWESSPSSNGFSLISK  
SHLWPOSKKQGIELICTGKINNNPRVAVPCGAPRPAQAARHGYSGRAGTARYCS  
RPCRAVPATGSAVPAARLARVPVPLWLSHQCAAPQPSLTMTFNWVRNVCRTTGP  
ROIICDOSVIGLSPVTS"

join(29524..29612,29887..29965,30113..30169,30207..30533)

/gene="P0463F06.8"

join(29524..29612,29887..29965,30113..30169,30207..30533)

/gene="P0463F06.8"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB17121.1"

/db\_xref="GI:11034597"

/translation="MKLIJHPTLMDKYPLHAVVPRERQPLSLOSTLSDFHOSSTIVP  
EKEKHILILRPSQLLHSRRRPRPTSSGKITRNLALLHAVIEVNSDALQOPA  
RADTEHASGHSGYCPGLHGVPILKDNIVKRDRLNTAGALLGVSVCVVDGAVTAGLRA  
AAAAAILILSKANPVNWSARR"

join(34905..34955,35683..36040,36863..36942)

/gene="P0463F06.9"

join(34905..34955,35683..36040,36863..36942)

/gene="P0463F06.9"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB17122.1"

/db\_xref="GI:11034598"







```
/gene="P0698G03.3"
complement(join(6926..8061,8244..8395,8496..9262))
/codon_start=1
/product="putative receptor kinase"
/protein_id="BABI7323.1"
/db_xref="GI:1107198"
/translaton="MVLPGRFRSTALQVSVLCVLRVFPVDPDIEWLIHGIPACP
PFTGHLKVASPFRRHRYEGVASYELCTDDKATQIDNGYFVYTGIDYSDFWV
VDNISDRNCLPWHKRIPIYVYDVSSDESSPINQIVELDPASRWSPFNCKV
EINNVPISVACMGTSVSFYVLTGRLSYIENLEPSCGYLAWTIPDWMGVRVAPAD
SDYEDVYFMKRGFRVQFPFRRHNRNFKDLMEISPKTGKGLSDWFSLSYTFN
MDCATIPSPNGLITRIILYLAIFMFWMLAPPCRYLLAPLVLIIFLAQYKLKRMITD
AVSKFLMQOGLPRTVAYTDITAITGHREKLGOGGYGVYKGLVLLPGDLHAKITL
NGYSNCGEEFIEVATIGRIHVNVRVLFVFCSEEMRRALVYEMPRGSLDKHIES
ERFSDKNEIALGARGINYLHQCQDMQILHFDIKPHNILLDDNFVPRVADFLAK
LYPRDSFVSDRALRGVGMVEMVSRFIIISGKSDVYSFGMLLEMLDWMGVRNADP
NADSSAKSPYGMWYDKLADQVDEISFNFMHLEKRLCLVGLWCIQMKSHDRPT
MSAEIENLEGGVDAQLVLPFPFCDDGMPPPQVMDYSFHSLSLSAISEDDGIAELA
SS"
join(9938..10001,10770..11578)
/gene="P0698G03.4"
join(9938..10001,10770..11578)
/gene="P0698G03.4"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BABI7324.1"
/db_xref="GI:1107197"
/translaton="MICLNYYVYISLVFLNYKSKFTPPRIIPHPAPFARLNEGGII
PRDSLFRPFSLSLGSFVRIAPAILADYINGARELGRFPFSPAFSPFRYKIYP
RASLFRPFSPLALVGFVAPCPSSSAVAELRFAAVAPVLLPTRRHLHRLRVIA
DPVRSASPADRRSTVIVIPNRAAFLRSRRSRRCPCGVSRPPFPAPFLSALSJG
AAPRWSPAVTIGARAPPPVPGRAGEAAGRALAAAKWAPRSARASIVDRADAW
APVPDVPAPSLG"
join(14575..14697,15133..15179,17078..17656,17671..18037,
18110..18307)
/gene="P0698G03.5"
join(14575..14697,15133..15179,17078..17656,17671..18037,
18110..18307)
/gene="P0698G03.5"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BABI7325.1"
/db_xref="GI:1107198"
/translaton="MESPTWGTFGDVPEGEVSALLIISPRGRVAPAPAPRAERTA
TFPRACPRKWSGRGSRADRVHKGKWPVSPVHVGFGARGWRRSAPWPVRSRTG
RGMTAGPTWQPRRRRRRRRRRRRAMPVTAGDHRGGAAPERRRGER
RSSAHPGTTTTTTTAAGVESGGAARDDDGAPTGVGRNGGADVDGDAAMPREVAP
SREEDRSDDGGEPERRKATARFGMPRATVLRHGAADGVGNTAKRMEVSPSRGDS
VGGEKVDGGEPGRGARREHSDGESERRAAETTERSTGSKYIATGGGGRGERKPR
AETSGATEDAGELRGVNPDESGERECKGGETGEGITGSGWRRRGRKAMTGSHP
ARVAGGPACORRARGEPHGRGREREREGELGRAGPREGGGGEF"
join(21957..22009,24616..25161,25209..25575,25648..25845)
/gene="P0698G03.6"
join(21957..22009,24616..25161,25209..25575,25648..25845)
/gene="P0698G03.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BABI7326.1"
/db_xref="GI:1107198"
/translaton="MWHGMRLLGCWRCADLCRSGRADRVHKGKWPVSPVTHGGPG
ARSWRRAPVPRVRSRTRGRRLTAGPTWQPRARRRRRRRRRRRRRAPVMTAGD
HRRGGAPEARERGESERSSAHPCTTTTTRTAAGVESSGAARDDDGADGATVGR
NGADEVDGDAAMPREVAPREEDRSDDGGERKATARGAAGDGGGAKRNEVSP
SREGDRSDGGEKVDGDEGGARREHSDGESERRAAETTERSTGSKYIATGGGR
GRGERPRABISGIEDAGELRGVNPDESGERECKGGETGEGITGSGWRRRGRKA
MTGSHPSARVAGGPACORRARGELMGRGREREREGELGRAGPREGGGGEF"
join(26540..28903,28967..32695)
/gene="P0698G03.7"
join(26540..28903,28967..32695)
/gene="P0698G03.7"
/codon_start=1
/product="putative polyprotein"
/protein_id="BABI7327.1"
/db_xref="GI:1107198"
/translaton="MAEQAKVHDLSPSVSGDDGEPNRRRARTPPPHQNPQKGEAL
ERVEKSAAPVAGGERRRGLLVYDGTSPQALQAVGALRRPPVVPVDPESA
ORWLDDVANLVMTAQALGAGGSAATTKTGAATTSASSRRRRRAAAVAHSAAPG
SSAPTRDRHGEDARDLIERRRNRRTPRATEGASSRSVSRQREDQSPMPAGG
VGCRAVASLRNWRPFRPTITEKYDGSVNPTEFLQVYVTTGTEAAGDQDRVMAF
PMALGQARLWMLNLPASVHSMEDLCQFTFMFGYTPRGEADLHAGVQDESL
RSYLOFCOVRTNTPICIPAHAVIYAFRGVPHNMLEKIASKEPQTTAEUPOLADRV
KREBATTNKSAGCSAASAAPGSAQGTGRDRRRKRSVSDGSHGLAVGVPRNTR
KGRPASDKAGTSPSHERPASKWCVSHTNSLHDLACRAVKNAIRKWEEDRROR
RREKSAVPSGKRSEAKQKAPAVDDDDDLGFGPEGATITAVDGGACAHVSHRS
FKAMRELLAAATTHEATRMRSEVALTIDQDHPCPVARGGQIAMVVSPTICNKL
GRVLIDGAAALNLSPAFDAIKAPGVLPLOPIGVTPGHTPLGHIDILPVTFGS
ANETRVNFDVADLSLPYNNALGRPALVKFMAAHVAYLQMKWPGGGPIISIGDLK
VALACMRADHAAATPEGGDRGLTSAPTRQRIATCDEPVVEEDALVSFLRA
FTVFAWRPADMPGVPREVIEHLAVRGPARGVQKVRQAPERQAFIREVAHLEA
GTFREYTHPELWLVVYVYKANGKLMCIDYDLNKPDKPDYPLPRIDQIVDTAGC
DLCLFLDAYSGYHOIRMARDEEKTFTTPVGYCYTSMPEGLKAGTFTQFTTRISL
GSOIGRNVAAVDDLVVVKTRNOETLLSDLAETFEENLSARKIKLPNDKCVGVPAKLL
GFLVSARGIEANPEKIRAIERMPPSKLRDQCVTCGCAALSRIISLGEKALPLFKL
LKRSQPTTEAEHALTQLKAYLSSPVLVAPNEPLLLYLAFTQVVSAAALVVER
DENNPHFVHSVLTWPGSERGGEAPSDGLRLPTTVGVPLPACQTVLGPADPOEG
KATAGRPHLTPDPEANPLTRPGKEGEEAPENGGORPLTTGVPLPACPTRRVQRP
DPDQPEATGRPLSSSDPEVIGTEKCAPKHLNERPGDTAPSKEDRPHRRVQRP
VYFVSEALRDAKTRYPOAKMLYAILMASRKLHYFOAHRVTVTSYPLGOILYNREG
TGRVVKAIELSEFDLHGFPHAIKSQLADFAVETPAPETVSIPEASNTSSQLPHT
THWVQFDGSLSGAGVLTSTQDVLRYLVLRDTRATNNAEYEGLAGURAA
GLIRLUULJGDSQVNVNVCYKCSQDQMDAYHVQRRMERHFDGTELHVRDRDN
MIADLSRLASSQVQPPGAFERLTQPSARPDLGETADPRPVRPVGLASGEGS
APSLRLIAWSEITQVLTDKTLPEDEGRESVQRIKSRVYLVLEGTLYRRAANGILIK
CIPQEGVELLADIHEGCGAHSRSLVGAERQGVWPTALADVLRRCRACQF
HAKQIHPAQALQIIPISWFPVAVGLDILGPFKAPGGFEYLYIAIDKFTKWPAYPV
VKIDHSALAFIKGTARFVGNRIITDNGQFTSELSFGDCEDMGKLCFASPFEFLY
YGEAVLPSELTLSPRATMYCEADQDLRRDDLDLEERRRRAALRAARYQQSLRRY
HORHVRASLVCVDDLVLRVQTRAGLSKLSPMWEGPYRVIGVPRPGSVRLATGDTGL
PNPNWIEHLRRFP"
complement(join(33311..34366,34450..35245,35345..36603))
alignment_scores:
Quality: 74.00 Length: 79
Ratio: 1.542 Gaps: 5
Percent Similarity: 60.759 Percent Identity: 34.177
alignment_block:
US-09-528-682-4 x AP002747
Align seg 1/1 to: AP002747 from: 1 to: 147857
10 AlaTyrGluProfile...AlaAsnThrAsnThrThrThrGlnPheLeuAs 25
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50054 GCTTTTGGCACCATTATAGATCAGATCAGGATCAGTACGAAATTTCTCGC 50103
25 nMetGlyAsnGluValAlaLeu.....AspGlyArgTyrSerAsnTyrA 40
||| ||||| ||||| ||||| ||||| ||||| |||||
50104 CATGATATCTCTGCTGCTCTCTCTCCTCAGCCATGCCACCTACAG...TACA 50150
40 laLeuTleSerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAla 56
||| ||||| ||||| ||||| ||||| ||||| |||||
50151 GTGCCACCACAGCTTCAGCTGCGGAGCAACAGACTTCTTCAGGCATTGC 50200
57 AsnIle.....AspGlyPheProGluValArgGluPheAsnSe 69
||| ||||| ||||| ||||| ||||| ||||| |||||
50201 CCAGTGTCCCGGTGCAGCAGCGGTGCAGCAGATCCCG..... 50239
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSer 81
||||| ||||| ||||| ||||| ||||| ||||| |||||
50240 .TTCCCTCAGCGCTCANTCCAGCAACACATCATCA 50275
seq_name: gb_htg:AC027163
seq_documentation_block:
```



```

LOCUS       AC027163             156690 bp      DNA      linear      HTG 04-OCT-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-607F22 map 15, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
ACCESSION   AC027163
VERSION     AC027163.2  GI:10567962
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 156690)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
             Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bida,F.,
             Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G.,
             Campopiano,A., Castle,A., Choquel,Y., Collangelo,M., Collins,S.,
             Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
             Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
             Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
             Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
             Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
             Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
             Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
             McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
             Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
             Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
             O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
             Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
             Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
             Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
             Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
             Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
             Young,G., Zainoun,J., Zimmer,A. and Zody,M.
             Direct Submission
             Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
             On Oct 4, 2000 this sequence version replaced gi:7331476.
             All repeats were identified using RepeatMasker:
             Smit, A.F.A. & Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             ----- Genome Center
             Center: Whitehead Institute/ MIT Center for Genome Research
             Center code: WIBR
             Web site: http://www-seq.wi.mit.edu
             Contact: sequence_submissions@genome.wi.mit.edu
             ----- Project Information
             Center project name: L6511
             Center clone name: 607_F_22
             ----- Summary Statistics
             Sequencing vector: M13; M77815; 100% of reads
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 0.960731
             Consensus quality: 150593 bases at least Q40
             Consensus quality: 153554 bases at least Q30
             Consensus quality: 154789 bases at least Q20
             Insert size: 153000; agarose-fp
             Insert size: 155490; sum-of-contigs
             Quality coverage: 5.1 in Q20 bases; agarose-fp
             Quality coverage: 5.0 in Q20 bases; sum-of-contigs
             -----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 13 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             * 1 1720: contig of 1720 bp in length
             * 1721 1820: gap of 100 bp

```

```

* 1821 4662: contig of 2842 bp in length
* 4663 4762: gap of 100 bp
* 4763 8750: contig of 3988 bp in length
* 8751 8850: gap of 100 bp
* 8851 15813: contig of 6963 bp in length
* 15814 15913: gap of 100 bp
* 15914 23306: contig of 7393 bp in length
* 23307 23406: gap of 100 bp
* 23407 33170: contig of 9764 bp in length
* 33171 33270: gap of 100 bp
* 33271 44422: contig of 11152 bp in length
* 44423 44522: gap of 100 bp
* 44523 81260: contig of 36738 bp in length
* 81261 81360: gap of 100 bp
* 81361 95970: contig of 14610 bp in length
* 95971 96070: gap of 100 bp
* 96071 106973: contig of 10903 bp in length
* 106974 107073: gap of 100 bp
* 107074 122796: contig of 15723 bp in length
* 122797 122896: gap of 100 bp
* 122897 144366: contig of 21470 bp in length
* 144367 144467: gap of 100 bp
* 144467 156690: contig of 12224 bp in length.

FEATURES             Location/Qualifiers
     source            1..156690
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
     misc_feature       1..1720
                     /clone="RP11-607F22"
                     /clone_lib="RPC1-11 Human Male BAC"
                     /note="assembly_fragment"
                     clone_end:SP6
                     vector_side:left
     misc_feature       1821..4662
                     /note="assembly_fragment"
     misc_feature       4763..8750
                     /note="assembly_fragment"
     misc_feature       8851..15813
                     /note="assembly_fragment"
     misc_feature       15914..23306
                     /note="assembly_fragment"
     misc_feature       23407..33170
                     /note="assembly_fragment"
     misc_feature       33271..44422
                     /note="assembly_fragment"
     misc_feature       44523..81260
                     /note="assembly_fragment"
     misc_feature       81361..95970
                     /note="assembly_fragment"
     misc_feature       96071..106973
                     /note="assembly_fragment"
     misc_feature       107074..122796
                     /note="assembly_fragment"
     misc_feature       122897..144366
                     /note="assembly_fragment"
     misc_feature       144467..156690
                     /note="assembly_fragment"
                     clone_end:17
                     vector_side:right

BASE COUNT  47316 a 29721 c 29869 g 48582 t 1202 others
ORIGIN

```

```

alignment_scores:
    Quality: 74.00      Length: 65
    Ratio: 1.805       Gaps: 1
    Percent similarity: 63.077  Percent Identity: 32.308

alignment_block:
    US-09-528-682-4 x AC027163/rev ..

```











```

sig_peptide      SAQITVNCA
                  /function="putative"
mat_peptide      join(847..1124,1180..1582)
                  /function="putative"
                  /product="unnamed"
intron           391 a 518 c 368 g 346 t
BASE COUNT
ORIGIN
alignment_scores
  Quality: 73.50      Length: 116
  Ratio: 1.065        Gaps: 4
Percent Similarity: 59.483 Percent Identity: 21.552
alignment_block:
US-09-528-682-4 x APEI4BXYL ..
Align seg 1/1 to: APEI4BXYL from: 1 to: 1623
1 AsptyrPheThrValAlaIleGlnAspLayrGluProIleAlaAsnTh 17
  :::::::::::::::::::: ::::::::::::::::::::
1262 GAGTACTAGTAGTCGAGAACTTCGGCACCTACGACCCCTCTCCAGGC 1311
  :::::::::::::::::::: ::::::::::::::::::::
17 rAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspG 34
  :::::::::::::::::::: ::::::::::::::::::::
1312 C.....ACCGTCAGGGCTCCGTGACCGCCGACG 1340
  :::::::::::::::::::: ::::::::::::::::::::
34 lyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAspArg 50
  :::::::::::::::::::: ::::::::::::::::::::
1341 GC.....TCCTCGTACAAAGATTGCACAGACCAG..... 1369
  :::::::::::::::::::: ::::::::::::::::::::
51 AspteuPheGlySerAlaAsnIleaspGlyPheProGluValArgGluPh 67
  :::::::::::::::::::: ::::::::::::::::::::
1370 ...CGCACCAACCGCCCTCCATCGACGGCACCCACAGACCTTCCAGCAGTA 1416
  :::::::::::::::::::: ::::::::::::::::::::
67 eAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnL 84
  :::::::::::::::::::: ::::::::::::::::::::
1417 CTGGTCTGTGGCCAGACAGCAGCGCGGCTCCGTCCATCATGATGAAGA 1466
  :::::::::::::::::::: ::::::::::::::::::::
84 ysGlnHisAlaAspPheLysTyrIleLysLeu..... 96
  :::::::::::::::::::: ::::::::::::::::::::
1467 CTCACCTTCGACGCGCTGGCTGCACAGGCGATGAAGTCGGAAGCAGCAAC 1516
  :::::::::::::::::::: ::::::::::::::::::::
97 .....LeuIleAsnAspGlyPhePheSerAsnAsnGlyGlyLys 110
  :::::::::::::::::::: ::::::::::::::::::::
1517 TACCAGATCGTGGCTACCGAGGGCTACTTCTCCAGCGCAGCGCTCAG 1564
  :::::::::::::::::::: ::::::::::::::::::::
seq_name: gb_htg:AC103259
seq_documentation_block:
LOCUS      AC103259      112484 bp      DNA      linear      HTG 21-DEC-2001
DEFINITION Rattus norvegicus clone CH230-166J4, *** SEQUENCING IN PROGRESS
            ***, 47 unordered pieces.
ACCESSION  AC103259
VERSION    AC103259.2 GI:17974771
KEYWORDS  HTG; HTGS_PHASE1.
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 112484)
AUTHORS  Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Banks,T., Barbara,J.,
            Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,

```

```

Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Monabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmari,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Woodden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 112484)
Worley,K.C.
Direct Submission
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:17062919.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJQF
Center clone name: CH230-166J4
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 101125 bases at least Q40
Consensus quality: 106675 bases at least Q30
Consensus quality: 112029 bases at least Q20
Estimated insert size: 98310; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 47 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
* 1 6637: contig of 6637 bp in length
* 6638 6737: gap of unknown length
* 6738 12170: contig of 5433 bp in length
* 12171 12270: gap of unknown length
* 12271 17400: contig of 5130 bp in length
* 17401 17500: gap of unknown length

```







Nucleopolyhedrovirus Genome  
 Virology 287 (2), 391-404 (2001)  
 MEDLINE 21425398  
 PUBMED 11531416  
 REFERENCE 2 (bases 1 to 139342)  
 Yu, J., Wang, L., Hu, X. and Pang, Y.  
 Direct Submission  
 Submitted (01-DEC-2000) State Key Laboratory for Biocontrol &  
 Institute of Entomology, Zhongshan University, 135 Xingang Road,  
 Guangzhou, Guangzhou 510275, China  
 Location/Qualifiers  
 1. 139342  
 /organism="Spodoptera litura nucleopolyhedrovirus"  
 /strain="G2"  
 /db\_xref="taxon:46242"  
 1. 750  
 /note="ORF1; major occlusion protein"  
 /codon\_start=1  
 /product="polyhedrin"  
 /protein\_id="AA01689.1"  
 /db\_xref="GI:15553211"  
 /translation="MYSRYSAVYVPHLGKTYVVDNKKYKNGHGVKNAKRRKHDALE  
 EADERELDHDKLVAEDEPMGKQKQLTLFEIRNVKPDTHKLYNNWNGKEFLRET  
 WTRFMEDESPVNDQVMDVFLVYNNRPTNRCFRFLAQLALRCDEYVPHDVIRIV  
 EPSYGTNNRYRISLAKGGGCPVMMHAEYTTSFESFDKVIWYFYKPIVYVGTDS  
 AEEELILEVSLVFKIKEFAPDAPLYGTAY"  
 complement(747..2393)  
 /note="ORF2; Acorf1629"  
 /codon\_start=1  
 /product="essential structural protein pp78/81"  
 /protein\_id="AA01690.1"  
 /db\_xref="GI:15553212"  
 /translation="MMAATLDELFLKANNYNVDLKLMSRLVTPSPCVELAKRKLISVSP  
 OSQVNLDIYECQLLRIDADDYKVSVPALSESLLKLIIDRAIDVQCDRQYE  
 LAMKHIDLAKESHDPQVKNHLERVLSLKLAVSTPPSPVGIENIDELTLTAATF  
 EPKSQEQEIEPPSKPEQSKSPIESILPELSKPIQKQSLKPIEPSILPEPP  
 IEPSPNEITIEPILPTQLPPPPPPQSIIMPPPPPPPPPPPPPPPPPPPPPPPP  
 QLITVPTTEARNLLLEQIKQSLRTEPAAPKPDYRNLLLEQIKQSLRTEPAAPK  
 QASLVNESPMAILSRIGIKPSSSSSEIETETPSSWLDDEEDPRRIKFKVSKSH  
 TOIGDENIGREIAREMYNRUKAETIASVIAESSEVIRNNAADILRLNNVKIV  
 AKDALLINVERTSIVQIMFEGDAESALAEQAGDQADTMSKDRFVRAVEHFIVN  
 DDIAKAKSLMQKGGDQDEKISQLRRRVDELRLSSQSI"  
 2392..3204  
 /note="ORF3; pk1"  
 /codon\_start=1  
 /product="protein kinase"  
 /protein\_id="AA01691.1"  
 /db\_xref="GI:15553213"  
 /translation="MDARFVEDINQCRDQVHSSCRLLDGKYGVSVIRHTPTDKLF  
 LKKIIRHENDALELTVHLMRDSHFVRLYAFNPKINVLIMDYVEDGDLFEILKT  
 DGPFSERKTAHVYSQLAFALNSLHGRVHNDVKNLVLTAKNRQVKLGDYGLKPV  
 GAVQKDGTVDSPEKIRDPYDETMDEWAGVIAVELTGEHPYKDGKDDLDVDT  
 LKRIKESFPRIQKRNISETALDFLKNLHPNIKSRICKKFKYIAVHDFLTKI"  
 complement(3519..5714)  
 /note="ORF4; hoar"  
 /codon\_start=1  
 /product="HOAR"  
 /protein\_id="AA01692.1"  
 /db\_xref="GI:15553214"  
 /translation="MUSNYIAFEESKMQCINFLVPANVKVQQRNNRVGHLDLOSN  
 TFRNFYTKYNTSVTLQVQKQKLRVPEAKLNKLNDLLDLTRFKIVEILLVKM  
 LKQKHENTYDSEYCLVRLFNATVDNTHKEITIAATGLVNMKRCAGQTVY  
 SHFLNCTTIRNEYIMKEKLPV/LNARQILEAKINLSINELKPVILLYOLRIDAIC  
 ESNKESASOLMTCGHRFCVDCIYEKIQEIKDPENYSCIICEKMYFISLNAAPNA  
 YMDFFYNQLRNATNANKIGITDDDDGVAATSDSTAKAKSKRSASTATNKS  
 TKRKLKLSLNSVSSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD  
 SDESDYADNDIPNPEAFVQKQKSGEQDQDQDQDQDQDQDQDQDQDQDQDQDQ  
 PEQRESSESAPE  
 DVIIKKIKISDPVIDQYKTKENQAAANENDDDDDDDDDDDDDDDDDDDDDDD  
 RDILKHLGFCLRLPCTNEFPVMEKVLIKDEKNEANGVPALFDSIDRYMCQARNI  
 HYTRERRRGIDENITNDELQQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ  
 AVSKSTAVSSIKSTAVSKSKSKYKQATISECNTTTLRLYPDGKVKFLTQ"  
 complement(6027..6230)

/note="ORF5"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA01823.1"  
 /db\_xref="GI:15553345"  
 /translation="MTVCASVLSPPDLINFFCSAKFGYRQDAIYTKIRIVDYLSS  
 KVTYVYIDSIIYIEIYIFALQKS"  
 complement(6227..6511)  
 /note="ORF6"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA01693.1"  
 /db\_xref="GI:15553215"  
 /translation="MVLITSTGAASKSTEVFEDYVKNKICFVRKLHLIYTHTPIKY  
 KLIKYLTFONPQNSNLYTWHTRASQTEFETESVVCQITNMFLLYRAR"  
 6447..8600  
 /note="ORF7"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA01694.1"  
 /db\_xref="GI:15553216"  
 /translation="MSSKTSVDFDAPVDVIVINKTITHPTSDISDYDAPEIFAGDOV  
 VASVHADATGCDVGCAGAAETAAGVAGADGAGDVGVAAGVADAAAGAAAGVGA  
 DAAAGAGNVILPVINISDKYNYNLYALDNCLNIIYEMELPDQIETIKTIVSN  
 EHLFSEYNTQTFLRCNVLDVNGRVYVNMNTDNYKIRIFCQYIVQIYLYIKNFDD  
 ANCDVSKLYSEIDFVELIRLDKMFKDFGLELYDMALRHTIMAKDNINKICFNLLIA  
 NAIGIRKSLVYMAFVNPYSSDAANHSDDMDYDNKNYSMTITRYIFDNHSHYDADMSKR  
 FNLDPEIYNSLOFEDYFAKENDYRDKINNCVFKLIIANVYISRYKSDNLYER  
 FNINISKAATSDDEVISIEKAARDLEKSESIHNSYSEILKSDSTFQTSNELLQNT  
 PFNPFOLDYALNDQCNKLSHLVSELSHNSIYNNVKNIYNIKOFNIMSREYFT  
 DTVLISKSRKIYHMFQVNMFCWIDTVYIINQVLDKIKIAQLTIRDEYDNDDEDDDD  
 DDSTINFDNINSNTEVSFOIEHVESPLKQPIVSELTPABESVSPQLPAADSEIES  
 KQTLIKLNPREIYANFLKKEAEVMAQAKQAEIMQTEAKETAEAVRAEVASSVSSKR  
 AKYNSVASKRRARIDSVSXKYGIKLPPRAAKRL"  
 8851..9720  
 /note="ORF8; ie-0; IE-0"  
 /codon\_start=1  
 /product="immediate early 0 protein"  
 /protein\_id="AA01695.1"  
 /db\_xref="GI:15553217"  
 /translation="MILHQSDNEINKPLSQVKHLRMSGVYDDTEMAIYAAATTTA  
 TDEKCNKNDVSSYHSDHQLVNSFLNMGYSDDLKYNLKAQYNKLAARFVHEKYNEDH  
 HDSGLDKTSPLYTGEDSVLGCDEKCYHDLIAETADLSKIEEDLYDASPVYRLNYFI  
 FVPYVRLQLQICMFTNDVCCKRSVRSTIARLETCLKRGNEKLESVRRLAKTLNMSV  
 FLEQNYVECNICKDVSNDREFLKPDECCGPRICNLCYANLWKESTNTPHPNCPVCKTS  
 YRVEKESSSPLI"  
 complement(9674..9829)  
 /note="ORF9"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA01824.1"  
 /db\_xref="GI:15553346"  
 /translation="MLITNCCSTEVSTTKLLSMIRVQNKYVLYILSRFIISVMMI  
 RTFQRDN"  
 9774..10268  
 /note="ORF10; dUTPase"  
 /codon\_start=1  
 /product="dUTP pyrophosphatase"  
 /protein\_id="AA01696.1"  
 /db\_xref="GI:15553218"  
 /translation="MDNNFVYETSEQQFVNNKQIESRPPOLKYLKSEHAYEPTYGS  
 DGAAGLDLKAHDYIIKNSRAVIAKTDLAICIPKDCYARIAPRSGLSVKYSIDVAGY  
 IDHDYRCNGVIGLNNGEFNVIIAETEWLTDLRAYLQMRARRGGIGFRNNTRRRREC  
 STGI"  
 10301..10800  
 /note="hrl; 4 copies of p-I repeat"  
 10801..12210  
 /note="ORF11"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA01697.1"  
 /db\_xref="GI:15553219"  
 /translation="MSSTTAPDALSLERNQLKYLASYFDLVDTQRLADETRPFIRE

repeat\_region  
 CDS







Consensus quality: 148585 bases at least Q40  
Consensus quality: 157081 bases at least Q30  
Consensus quality: 164382 bases at least Q20  
Estimated insert size: 152520; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 2.3x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 67 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
\* 8370: contig of 8370 bp in length  
\* 8371: gap of unknown length  
\* 8471: contig of 4870 bp in length  
\* 13341: gap of unknown length  
\* 13341: gap of unknown length  
\* 13441: contig of 8839 bp in length  
\* 22280: gap of unknown length  
\* 22379: contig of 7858 bp in length  
\* 30037: gap of unknown length  
\* 30137: contig of 5961 bp in length  
\* 30138: contig of 5961 bp in length  
\* 36099: gap of unknown length  
\* 36199: contig of 6640 bp in length  
\* 42839: gap of unknown length  
\* 42939: contig of 4341 bp in length  
\* 47280: gap of unknown length  
\* 47379: contig of 6076 bp in length  
\* 47380: contig of 6076 bp in length  
\* 53456: gap of unknown length  
\* 53555: gap of unknown length  
\* 53556: contig of 3034 bp in length  
\* 56590: gap of unknown length  
\* 56690: contig of 5927 bp in length  
\* 62617: gap of unknown length  
\* 62717: contig of 2880 bp in length  
\* 65597: gap of unknown length  
\* 65597: contig of 3233 bp in length  
\* 68930: gap of unknown length  
\* 69029: contig of 2846 bp in length  
\* 71876: gap of unknown length  
\* 71876: contig of 2719 bp in length  
\* 71976: gap of unknown length  
\* 74695: gap of unknown length  
\* 74795: contig of 3579 bp in length  
\* 78374: gap of unknown length  
\* 78474: contig of 3572 bp in length  
\* 82046: gap of unknown length  
\* 82146: contig of 3434 bp in length  
\* 85579: gap of unknown length  
\* 85580: contig of 2352 bp in length  
\* 88032: gap of unknown length  
\* 88132: contig of 2697 bp in length  
\* 90829: gap of unknown length  
\* 90929: contig of 2711 bp in length  
\* 93640: gap of unknown length  
\* 93739: contig of 2458 bp in length  
\* 96197: gap of unknown length  
\* 96297: contig of 3023 bp in length  
\* 96298: gap of unknown length  
\* 99321: contig of 1647 bp in length  
\* 101067: gap of unknown length  
\* 101167: contig of 3853 bp in length  
\* 101168: gap of unknown length  
\* 105021: contig of 2139 bp in length  
\* 105121: gap of unknown length  
\* 107259: contig of 2344 bp in length  
\* 107260: gap of unknown length  
\* 107360: contig of 2344 bp in length  
\* 109704: gap of unknown length  
\* 109804: contig of 3148 bp in length  
\* 112952: gap of unknown length  
\* 113052: contig of 2869 bp in length  
\* 115921: gap of unknown length  
\* 116020: contig of 2271 bp in length  
\* 118291: gap of unknown length  
\* 118391: contig of 3101 bp in length  
\* 121492: gap of unknown length  
\* 121493: contig of 2051 bp in length  
\* 121593: gap of unknown length  
\* 123644: contig of 2565 bp in length  
\* 123744: gap of unknown length  
\* 126309: contig of 1450 bp in length  
\* 126409: gap of unknown length  
\* 127859: contig of 2203 bp in length  
\* 127959: gap of unknown length  
\* 130162: contig of 2811 bp in length  
\* 130262: gap of unknown length  
\* 133073: contig of 3225 bp in length  
\* 133173: gap of unknown length  
\* 136398: contig of 2192 bp in length  
\* 136498: gap of unknown length  
\* 138690: contig of 1573 bp in length  
\* 138790: gap of unknown length  
\* 140363: contig of 1996 bp in length  
\* 140463: gap of unknown length  
\* 142459: contig of 1733 bp in length  
\* 142559: gap of unknown length  
\* 144292: contig of 1870 bp in length  
\* 144392: gap of unknown length  
\* 146262: contig of 2050 bp in length  
\* 146362: gap of unknown length  
\* 148412: contig of 2163 bp in length  
\* 148512: gap of unknown length  
\* 150675: contig of 1861 bp in length  
\* 150775: gap of unknown length  
\* 152636: contig of 1585 bp in length  
\* 152736: gap of unknown length  
\* 154321: contig of 1199 bp in length  
\* 154421: gap of unknown length  
\* 155620: contig of 2381 bp in length  
\* 155720: gap of unknown length  
\* 158101: contig of 2291 bp in length  
\* 158201: gap of unknown length  
\* 160491: contig of 1633 bp in length  
\* 160592: gap of unknown length  
\* 162224: contig of 1083 bp in length  
\* 162325: gap of unknown length  
\* 163408: contig of 1802 bp in length  
\* 163508: gap of unknown length  
\* 165310: contig of 1426 bp in length  
\* 165410: gap of unknown length  
\* 166836: contig of 1277 bp in length  
\* 166936: gap of unknown length  
\* 168213: contig of 1628 bp in length  
\* 168313: gap of unknown length  
\* 169941: contig of 1699 bp in length  
\* 170041: gap of unknown length  
\* 171740: contig of 1012 bp in length  
\* 171840: gap of unknown length  
\* 172852: contig of 1012 bp in length

## alignment\_scores:

Quality: 73.50 Length: 67  
Ratio: 1.670 Gaps: 2  
Percent Similarity: 65.672 Percent Identity: 32.836

## alignment\_block:

US-09-528-682-4 x AC095438/rev ..

Align seg 1/1 to reverse of: AC095438 from: 1 to: 187586

41 Leu1leSerAlaGluGlyMetAspArgAspLeuPheGlySerAlaAs 57  
||| ||||| ||| :|||  
68178 TTAACTCTGACACTCGGTGAGCTCAAGGGCAATCTGTCTACAGATGAG 68129  
57 nileAspGlyPheProGluValArgGluPheAsnSerLeuProAsnL 74  
||| ||| ||||| :|||  
|||||











```
/gene="nifH"  
/codon_start=1  
/transl_table=11  
/product="dinitrogenase reductase"  
/protein_id="AAC45512.1"  
/db_xref="GI:1666883"  
/translation="MVRKIAIYGGKIGKSTTQNTVAAMAHFHDKKVFIHGCDPKAD  
STRLLHGQVQVMTDTRERKGEDECTPKVIEVFGGVKCVESGGPEPGVGCAGRV  
ITAITLMSQHGVEDLDFVFDVLGVVCGGFAMPVRDRADEIYVIVASGEMMALYA  
ANNICKGMVKVAEQSGVRLGGIICNSRVNDELDDLOEFCDKIGTQLIHFPRDNIVQ  
RAFOKKAVVDVDDTCNOALEYKELARKI IENENLVIPTMTDELELT SKYGFLD"  
998..1002  
/gene="nifII"  
998..1326  
/gene="nifII"  
/note="glnBI"  
1009..1326  
/gene="nifII"  
/note="similar to PII"  
/codon_start=1  
/transl_table=11  
/product="GlnBI"  
/protein_id="AAC45513.1"  
/db_xref="GI:1809124"  
/translation="MKMIRAVVRPSKAEEVVDALAESGCLATKMDVIGRGKQKGIKI  
DQIYYDELPTMLVVEDDTAENVIELITKTAYTSGFYGRIFVSPVDEAYTVTRTS  
CGI"  
1027..1701  
/gene="nifI2"  
/note="glnBI1"  
1327..1330  
/gene="nifI2"  
1336..1701  
/gene="nifI2"  
/note="similar to PII"  
/codon_start=1  
/transl_table=11  
/product="GlnBI1"  
/protein_id="AAC45514.1"  
/db_xref="GI:1809125"  
/translation="MKEIIAIIIPSKMAQTKVLECLGFPAMTANRVLGRGKOKAIVG  
ELGFVDNKELLNQPGDMRYIPKTLTLVDPEDASLVVLEALMKVNKSGQYGDGKIFV  
CPTEIDITVTRSERGEAT"  
1741..3184  
/gene="nifD"  
1741..1744  
/gene="nifD"  
1751..3184  
/gene="nifD"  
/codon_start=1  
/transl_table=11  
/product="dinitrogenase alpha subunit"  
/protein_id="AAC45515.1"  
/db_xref="GI:1666886"  
/translation="MPFCLLDVDKDIPEREQHIYIKDSKEPKCHQKQKQNTNTPGSM  
TERGCAGVKGVIITGAIKDVLHVHSPVCGTAYNGTTRKRYTPREMPDGSVPVEN  
FNLHIVGDTLDSVDVFGGMKLVIREASKEFPVNAIYVYATCTGLIGDDLDA  
VCKEMAEGLKDVAVNAGPAGTOSKGHVGNVYIFENLVGTKEPKTQTDYDNL  
GEYNIGDWDVLEKYFEDMGINVLSTGDTGTHGELCWMHAKLSIVRCORSATYVAK  
LIEEKYGPYLKVFDFGPEYCANLRAVKYFGKEIEAEAVIQKEMEKIQPELDIFYOS  
KLGQKIWISAGPQKSWHLSKPIEQYLGMDDVVALSGLFEHEDGFKMQRKADGTLLI  
DDPTLEMEVEVKYQPEIYVLGIGKEKIFFKLGVPSVMIHSENGPYTGFGGFVNNA  
RDIPTAIYNPAWKLGMFGEGEPGDSNE"  
3167..4565  
/gene="nifK"  
3167..3170  
/gene="nifK"  
3177..4565  
/gene="nifK"  
/codon_start=1  
/transl_table=11  
/product="dinitrogenase beta subunit"  
/protein_id="AAC45516.1"
```

```
/db_xref="GI:1666887"  
/translation="MSELNVITKERTAVINPIVTCQPLGAMVYSGIERGMPLVHGSO  
GCSTFVRYGFARHFREPADIATVSLHEDAAVFGKKNLISGLNLAARFKPDVGMVYT  
TCSSEIIGDDVAGFIKTAKVETAKMGEEAAKIKI VQIOTPSFVEHOFKGYDNAIKR  
IVDTLAEPKDEENGKLI IIPGIVNPGDIRKHMLSLMVEGILLITDTSDFDPLRP  
SKADNPYQKGGTPLADQDCANSLGTISLANVANSAPASLEKYNMPSKVSAPIG  
IQNDSFRTVYKGTGNDVTDIILDERGIVIDAMADVASRYLFGKKVAIYGDPSITVG  
MARVAELGMIPIKVYCTGVKNIEYFVNDLKKVAKSEDEIDALFGQDLRALDVLKENP  
VDLMIGSDSDGIMAKDLGIPLYRVGPVYDRVQORRPIIGTNGALNLDVGTITNLD  
KYYETQDWKLOQ"  
4578..6039  
/gene="nifE"  
4578..4580  
/gene="nifE"  
4588..6039  
/gene="nifE"  
/codon_start=1  
/transl_table=11  
/product="NifE"  
/protein_id="AAC45517.1"  
/db_xref="GI:1666888"  
/translation="MVLNLDTENRRKMDGNDNDDDFDLEVPINSPIFEKLSIEALKAR  
ESPMCVSGKDDSIPTCDQNSTPGMITQRCVYGGARVYLMPTTDAVHLVHGPVIGCAAC  
TWDIRGSKTGDKLYKNGFSTDLQEKDIVFGGKLYESILEVKNLYHPGALFYVSTC  
VYGLIGDDLKAVCRQAQATGCRVIPVQSEGFKSNTAGHKLACDMLDYYVIGTEEP  
EEHPYSINIGEFNVAGDLWGLIIPLYEKMGVYHTAITGDSVAKVASAHSKLNIV  
QCPSKSNYLAQMDDKKGIPSPKVNFFGLDETFSKRAVAEFEGDEEMIKRTEELIKS  
EIKNLRDEITSEYQKDLSGRTVAIYSGAHSKSWALSAFGEIDMEIIMSGTQNKPKEDIQ  
QIRHVCETGLIVDDASSMELVQLLKEYKPDILLISGAKEKYLKSGIPHCDFNHDRI  
TAFSGYQGFNFARVVHTAVMTPIWRLSRKMMI"  
6039..7425  
/gene="nifN"  
6039..6042  
/gene="nifN"  
6049..7425  
/gene="nifN"  
/codon_start=1  
/transl_table=11  
/product="NifN"  
/protein_id="AAC45518.1"  
/db_xref="GI:1666889"  
/translation="MTEKKFVAINPCKICQPMGAIFALMGVKNKSMPLIHSGHGCSTYM  
RQLGRHPEAVNVASSSSSEANAVLGGEENLKHAI RNYSKIYGPPIIGVTTSCLSSET  
IGDVEGIIIRLFKQNDLPEIVPSTPSYKTHVGEYDSTVKALVENLATPSKPSK  
LNVTGMVSPGDIATQIRILREMGIESIITDTSMDAPDFGSEVSVSSAGTVVEEL  
KDTANSEATISLCKHANSAAKFLHEAYGINSVSPYLPGLQSTDFEPMELCKLTNKEI  
PEQITRERKGLIDAMIDAHQHNFGKVAIYGDVVVGFREFVLELGMIPSIIVCTGSE  
SKTSVGDINAVTKDSKHPVILAPGLDYLHQBIEIKKSKVDLLIGNSYGSRIAGAENIP  
LIRMGEPYDRVGAQRITSLGYTAGIRLVDTITNTILDYFDFEGWALNENTEVDHGE  
FESSNCIN"  
7381..7711  
/gene="nifX"  
7381..7384  
/gene="nifX"  
7391..7711  
/gene="nifX"  
/codon_start=1  
/transl_table=11  
/product="NifX"  
/protein_id="AAC45519.1"  
/db_xref="GI:1666890"  
/translation="MESLKVATSTDGVNVMHFGAATHFFIFEIKDNVAEPLEFRGN  
PKRHKDHDTRWNVALELLKDCGAI FCSRMGDPKSVLEKGMKVVMTGTNLDALKE  
FLTA"  
BASE COUNT 2911 a 1385 c 1625 g 2034 t  
ORIGIN  
  
alignment_scores:  
Quality: 72.50 Length: 85  
Ratio: 1.394 Gaps: 3  
Percent Similarity: 61.176 Percent Identity: 27.059
```







61 he.....ProGluValArgLysPhe.AsnSerLeu.ProAsnAsnL 74  
||| ||||| :||||| | |||||  
13931 TTGTCCAGACCTTCAAGCGCATTATGAGTCTGAAGAACCACA 13882

74 ysaAlaserSer.....AspThrAlaserLeuAsnLynHisasp 87  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
13881 GTGCTACTGGTGCGGCCCTTAAGAGAACAGCTTAGATTACAACACCAC 13832

88 AlaAspPhelyLysTyrlleLysLeuLlleAsnAsnAspGlyphe 103  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
13831 AGTTCCTTC...AAATCATCGAAGACCTTCCTCAATAATGATGGCTAC 13787

```
seq_name: gb_htg.AC099184

seq_documentation_block:
LOCUS AC099184          166852 bp    DNA        linear   HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-117A22, *** SEQUENCING IN PROGRESS
***, 55 unordered pieces.
ACCESSION AC099184
VERSION AC099184.2 GI:17973062
KEYWORDS HTG; HTGS-PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 166852)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaraturung,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Blinag,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.I., Ding,X., Dinh,H.H., Douthevaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseg,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokoko,S.,
Oguh,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs.R.
Direct Submission
Unpublished
2 (bases 1 to 166852)
Worley,K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
```

Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:16874678.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GHAF  
Center clone name: CH230-117A22  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329First call to findPhrapList  
Consensus quality: 140999 bases at least Q40  
Consensus quality: 148978 bases at least Q30  
Consensus quality: 156080 bases at least Q20  
Estimated insert size: 147801: sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft.data.html).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 55 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1	7805:	contig of 7805 bp in length
*	7806	7905: gap of unknown length
*	7906	17053: contig of 9148 bp in length
*	17054	17153: gap of unknown length
*	17154	24346: contig of 7193 bp in length
*	24347	24446: gap of unknown length
*	24447	30009: contig of 5563 bp in length
*	30010	30109: gap of unknown length
*	30110	37096: contig of 6987 bp in length
*	37097	37196: gap of unknown length
*	37197	41979: contig of 4783 bp in length
*	41980	42079: gap of unknown length
*	42080	47598: contig of 5519 bp in length
*	47599	47698: gap of unknown length
*	47699	51706: contig of 4008 bp in length
*	51707	51806: gap of unknown length
*	51807	55084: contig of 3278 bp in length
*	55085	55184: gap of unknown length
*	55185	59280: contig of 4096 bp in length
*	59281	59380: gap of unknown length
*	59381	63482: contig of 4102 bp in length
*	63483	63582: gap of unknown length
*	63583	66447: contig of 2865 bp in length
*	63448	66547: gap of unknown length
*	63548	68653: contig of 2106 bp in length
*	63654	68753: gap of unknown length
*	63754	73787: contig of 5034 bp in length
*	73788	73887: gap of unknown length
*	73888	76434: contig of 2547 bp in length
*	76435	76534: gap of unknown length
*	76535	79717: contig of 3183 bp in length
*	79718	79817: gap of unknown length
*	79818	82511: contig of 2694 bp in length
*	82512	82611: gap of unknown length
*	82612	85950: contig of 3339 bp in length
*	85951	86050: gap of unknown length
*	86051	90186: contig of 4136 bp in length
*	90187	90286: gap of unknown length
*	90287	92569: contig of 2383 bp in length
*	92670	92769: gap of unknown length
*	92770	95162: contig of 2393 bp in length
*	95163	95262: gap of unknown length
*	95263	97444: contig of 2182 bp in length







Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C.,  
 Hollins, B., Honsi, F., Howard, S., Huber, J., Huly, S., Hume, J.,  
 Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C.,  
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
 Loulsegod, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawhiney, E., McLeod, M. P., Meador, M.,  
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S.,  
 Ogulu, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,  
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N.,  
 Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

Unpublished  
 2 (bases 1 to 172199)

Worley, K. C.

Direct Submission

Submitted (15-NOV-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:16930859.

COMMENT

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GIID  
 Center clone name: CH230-62L13  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 138342 bases at least Q40  
 Consensus quality: 145173 bases at least Q30  
 Consensus quality: 150824 bases at least Q20  
 Estimated insert size: 142168; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-gel estimation  
 Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 63 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 7563: contig of 7563 bp in length  
 \* 7564 7663: gap of unknown length  
 \* 7664 17066: contig of 9403 bp in length  
 \* 17067 17166: gap of unknown length  
 \* 17167 23634: contig of 6488 bp in length  
 \* 23635 23754: gap of unknown length  
 \* 23755 31080: contig of 7326 bp in length  
 \* 31081 31180: gap of unknown length  
 \* 31181 36505: contig of 5325 bp in length  
 \* 36506 36605: gap of unknown length  
 \* 36606 40235: contig of 3630 bp in length  
 \* 40236 40335: gap of unknown length

40336 44553: contig of 4218 bp in length  
 44554 44653: gap of unknown length  
 44654 47729: contig of 3076 bp in length  
 47730 47829: gap of unknown length  
 47830 52693: contig of 4864 bp in length  
 52694 52793: gap of unknown length  
 52794 56955: contig of 4162 bp in length  
 56956 61317: contig of 4262 bp in length  
 61318 61417: gap of unknown length  
 61418 64575: contig of 3158 bp in length  
 64576 64675: gap of unknown length  
 64676 66371: contig of 1696 bp in length  
 66372 66471: gap of unknown length  
 66472 68616: contig of 2145 bp in length  
 68617 68716: gap of unknown length  
 68717 72927: contig of 4211 bp in length  
 72928 73027: gap of unknown length  
 73029 76039: contig of 3012 bp in length  
 76040 76139: gap of unknown length  
 76140 78427: contig of 2288 bp in length  
 78428 78527: gap of unknown length  
 78528 81683: contig of 3156 bp in length  
 81684 81783: gap of unknown length  
 81784 83786: contig of 2003 bp in length  
 83787 83886: gap of unknown length  
 83887 86932: gap of unknown length  
 86933 89492: contig of 2560 bp in length  
 89493 89592: gap of unknown length  
 89593 92294: contig of 2702 bp in length  
 92295 94515: contig of 2121 bp in length  
 94516 94615: gap of unknown length  
 94616 98144: contig of 3529 bp in length  
 98145 98244: gap of unknown length  
 98245 101425: contig of 3181 bp in length  
 101426 101525: gap of unknown length  
 101526 103718: contig of 2193 bp in length  
 103719 103818: gap of unknown length  
 103819 105824: contig of 2006 bp in length  
 105825 105924: gap of unknown length  
 105925 107976: contig of 2052 bp in length  
 107977 108077: contig of 2533 bp in length  
 108078 110610: gap of unknown length  
 110611 113479: contig of 2770 bp in length  
 113480 113579: gap of unknown length  
 113580 115824: contig of 2245 bp in length  
 115825 115924: gap of unknown length  
 115925 118129: contig of 2205 bp in length  
 118130 118229: gap of unknown length  
 118230 119998: contig of 1769 bp in length  
 119999 120099: gap of unknown length  
 120100 122191: contig of 2093 bp in length  
 122192 122291: gap of unknown length  
 122292 125454: contig of 3163 bp in length  
 125455 125554: gap of unknown length  
 125555 126955: contig of 1401 bp in length  
 126956 127055: gap of unknown length  
 127056 128648: contig of 1593 bp in length  
 128649 128748: gap of unknown length  
 128749 131298: contig of 2550 bp in length  
 131299 131398: gap of unknown length  
 131399 134230: contig of 2832 bp in length  
 134231 134330: gap of unknown length  
 134331 136037: contig of 1707 bp in length  
 136038 136137: gap of unknown length  
 136138 138225: contig of 2088 bp in length  
 138226 138325: gap of unknown length  
 138326 140301: contig of 1976 bp in length  
 140302 142114: contig of 1713 bp in length  
 142115 140402



```

* 142115 142214: gap of unknown length
* 142215 143450: contig of 1236 bp in length
* 143451 143550: gap of unknown length
* 143551 145712: contig of 2162 bp in length
* 145713 145812: gap of unknown length
* 145813 147823: contig of 2011 bp in length
* 147824 147923: gap of unknown length
* 147924 149107: contig of 1184 bp in length
* 149108 149207: gap of unknown length
* 149208 151034: contig of 1827 bp in length
* 151035 151134: gap of unknown length
* 151135 152240: contig of 1106 bp in length
* 152241 152340: gap of unknown length
* 152341 154103: contig of 1763 bp in length
* 154104 154203: gap of unknown length
* 154204 155590: contig of 1387 bp in length
* 155591 155690: gap of unknown length
* 155691 156842: contig of 1152 bp in length
* 156843 156942: gap of unknown length
* 156943 158009: contig of 1067 bp in length
* 158010 158109: gap of unknown length
* 158110 159685: contig of 1576 bp in length
* 159686 159785: gap of unknown length
* 159786 160998: contig of 1213 bp in length
* 160999 161098: gap of unknown length
* 161099 162572: contig of 1474 bp in length
* 162573 162672: gap of unknown length

alignment_scores:
    Quality: 72.50      Length: 155
    Ratio: 1.098       Gaps: 5
    Percent Similarity: 42.581      Percent Identity: 21.290

alignment_block:
US-09-528-682-4 x AC099452/rev ..

Align seg 1/1 to reverse of: AC099452 from: 1 to: 172199

6  ArgileGlnAspAlaTyrgluProIleAlaAsnThrAsnThrThrcI 22
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29064 AGAATCTCAAGCGCAGAGATACCATAGAAAACATCAACATCTGTGCAAA 29015
22 nPheLeu.....AsnMetGlyAsnGluValAlaL 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29014 GATAATGTAACACGCAAAAGTTCATAGCAACAAACACAGGAATACAGA 28965
32 euAspGlyArgTyrgSerAsnTyrrAlaLeuIleSerAlaGluGlyMet 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28964 ACACAATGAGAAGATCAACACCTTAAGGATAGTAAAGTATAGAGAAAGTGAA 28915
49 AspArgAspLeuPheGlySerAlaAsnIle.....As 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28914 GAATCCCACTTAAGGGCCGCGCATTAATATGTTCAACAAAATATAGAGA 28865
59 pGlyPheProGluValArg..... 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28864 TAACTTCCCTAACCTAAAGAAAGAGATGCCAGAAACATAGAGAGCCT 28815
66 .....GluPhe..... 67
28814 ACAGAACTCCAATTAAGAACAGCGGAGCAGCTACCGCAATTCCTCTAAGA 28765
67 ..... 67
28764 AGTTCAATTAATTTATACCAAAACACACAAAGACCCCAACAAAGAGA 28715
68 .AsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnL 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28714 GAACITCAGACCAACTCTCCTTGTGNAGACTGACGCAAGAGATCAATCA 28665
84 ysGlnHisAspAlaAspPheLysLysTyrrIleLysLeuLeuIleAsnAsn 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28664 AAATTTCTTCAAACTGGATCCCAAGAACATGTCAAGCATCATCATCAT 28615

```

101 Asp.....GlyPhe 103

```

|||||
28614 GATCAAGTAGGCTTC 28600

```

seq\_name: gb\_htg:AC094036

seq\_documentation\_block:

LOCUS AC094036 182583 bp DNA linear HTG 21-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-34L18, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 71 unordered pieces.

ACCESSION AC094036

VERSION AC094036.2 GI:17975630

KEYWORDS HTG; HTGS, PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 182583)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Briefe,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Loisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,  
 Oguh,N., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuben,I., Rolfe,M.,  
 Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,  
 Sisson,I., Sodergren,E., Sonaikar,F., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalobos,C., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

TITLE Direct Submission

REFERENCE 2 (bases 1 to 182583)

Worley,K.C.

Direct Submission

AUTHORS

Submitted (14-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On Dec 21, 2001 this sequence version replaced gi:15617897.

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)



```
----- Project Information
Center project name: GPHK
Center clone name: CH230-34L18
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 140876 bases at least Q40
Consensus quality: 157668 bases at least Q30
Consensus quality: 167320 bases at least Q20
Estimated insert size: 163483; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hqsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 71 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
*
* 1
* 6528: contig of 6527 bp in length
* 6628: gap of unknown length
* 11290: contig of 4663 bp in length
* 11391: gap of unknown length
* 11391: contig of 3428 bp in length
* 14818: gap of unknown length
* 14819: contig of 4684 bp in length
* 19603: gap of unknown length
* 19702: contig of 4421 bp in length
* 19703: gap of unknown length
* 24124: contig of 4447 bp in length
* 24224: gap of unknown length
* 28670: contig of 4447 bp in length
* 28770: gap of unknown length
* 33216: contig of 4446 bp in length
* 33217: gap of unknown length
* 33316: gap of unknown length
* 33317: contig of 3491 bp in length
* 36807: gap of unknown length
* 36808: gap of unknown length
* 40610: contig of 3703 bp in length
* 40710: gap of unknown length
* 44497: contig of 3787 bp in length
* 44597: gap of unknown length
* 44598: contig of 3172 bp in length
* 47769: gap of unknown length
* 47770: gap of unknown length
* 47869: gap of unknown length
* 52630: contig of 4761 bp in length
* 52730: gap of unknown length
* 55734: contig of 3024 bp in length
* 55755: gap of unknown length
* 55854: gap of unknown length
* 60016: contig of 4162 bp in length
* 60116: gap of unknown length
* 60117: contig of 2594 bp in length
* 62710: contig of 2594 bp in length
* 62711: gap of unknown length
* 67066: contig of 4256 bp in length
* 67067: gap of unknown length
* 71934: contig of 4768 bp in length
* 71935: gap of unknown length
* 72034: gap of unknown length
* 72035: contig of 2550 bp in length
* 74584: gap of unknown length
* 74585: gap of unknown length
* 77692: contig of 3008 bp in length
* 77693: gap of unknown length
* 80291: contig of 2499 bp in length
* 80292: gap of unknown length
* 80391: gap of unknown length
* 83294: contig of 2903 bp in length
* 83295: gap of unknown length
* 83395: contig of 2458 bp in length
* 85852: gap of unknown length
* 85853: gap of unknown length
* 85952: contig of 2616 bp in length
* 88568: gap of unknown length
* 88569: gap of unknown length
* 91156: contig of 2488 bp in length
* 91157: gap of unknown length
* 94694: contig of 3438 bp in length
* 94794: gap of unknown length
* 94795: contig of 2492 bp in length
* 97286: gap of unknown length
* 97287: contig of 2796 bp in length
* 97387: gap of unknown length
* 100183: contig of 3534 bp in length
* 100282: gap of unknown length
* 100283: contig of 3534 bp in length
* 103917: gap of unknown length
* 106431: contig of 2515 bp in length
* 106432: gap of unknown length
* 106531: gap of unknown length
* 108431: contig of 1900 bp in length
* 108432: gap of unknown length
* 108531: contig of 1797 bp in length
* 108532: gap of unknown length
* 110328: contig of 1795 bp in length
* 110329: gap of unknown length
* 110429: contig of 1795 bp in length
* 112224: gap of unknown length
* 112323: contig of 2810 bp in length
* 115133: gap of unknown length
* 115134: contig of 2458 bp in length
* 117691: gap of unknown length
* 117692: contig of 2458 bp in length
* 117992: gap of unknown length
* 120331: contig of 2540 bp in length
* 120332: gap of unknown length
* 123409: contig of 2878 bp in length
* 123410: gap of unknown length
* 125725: contig of 2316 bp in length
* 125726: gap of unknown length
* 125825: contig of 1573 bp in length
* 127398: gap of unknown length
* 127498: contig of 1530 bp in length
* 129028: gap of unknown length
* 129128: contig of 1961 bp in length
* 129129: gap of unknown length
* 131090: contig of 2649 bp in length
* 131190: gap of unknown length
* 133838: contig of 1844 bp in length
* 133839: gap of unknown length
* 135783: contig of 1519 bp in length
* 135882: gap of unknown length
* 137401: contig of 1519 bp in length
* 137501: gap of unknown length
* 138927: contig of 1426 bp in length
* 139027: gap of unknown length
* 141126: contig of 2099 bp in length
* 141226: gap of unknown length
* 143512: contig of 2286 bp in length
* 143513: gap of unknown length
* 143612: contig of 1550 bp in length
* 145162: gap of unknown length
* 145262: contig of 1867 bp in length
* 147129: gap of unknown length
* 147229: contig of 1686 bp in length
* 148915: gap of unknown length
* 149015: contig of 1686 bp in length
* 150648: gap of unknown length
* 150748: contig of 1633 bp in length
* 150749: gap of unknown length
* 152071: contig of 1323 bp in length
* 152072: gap of unknown length
* 153947: contig of 1776 bp in length
* 154047: gap of unknown length
* 155822: contig of 1775 bp in length
* 155922: gap of unknown length
* 157858: contig of 1936 bp in length
* 157958: gap of unknown length
* 157959: contig of 1645 bp in length
* 159604: gap of unknown length
* 159704: contig of 1867 bp in length
* 161571: gap of unknown length
* 161670: gap of unknown length
```

```
alignment_scores:
  Quality: 72.50      Length: 122
  Ratio: 1.151       Gaps: 6
  Percent Similarity: 51.639  Percent Identity: 27.049
alignment_block:
  US-09-528-682-4 x AC094036
Align seg 1/1 to: AC094036 from: 1 to: 182583
```







```

61 he.....ProGluValArgLuphe.AsnSerLeu.ProAsnAsnL 74
|| |||::: :::::|| | |||| |:::|
7734 TTGTCCAAAGACCTTCAAGGCAGCATTCTATGAGTCTGTGAAGAACCACA 7685
74 ysAlaSerSer.....AspThrAlaSerLeuAsnLysLHisAsp 87
||||:|::: |:::|::: |:::|::: |:::|::: |:::|::: |:::|:::
7684 GTGCTACTGGGGTCCCCTAAGAAGAACAGCTTAGATTACACACCAA 7635

88 AlaAspPheLysTyRileLysLeuLeuIleAsnAsnAspGlyphe 103
:::::|::: |:::|:::|::: |:::|:::|:::|:::|:::|:::|:::|:::
7634 AGTTCTTC...AAATATCTGAAAAGCCTTCTCAATAATGAGCTAC 7590

seq_name: gb_pl:TAVDAC1

seq_documentation_block:
LOCUS TAVDAC1 1159 bp mRNA linear PLN 28-FEB-1996
DEFINITION T.aestivum VDAC 1 mRNA.
ACCESSION X77733
VERSION X77733.1 GI:456671
KEYWORDS VDAC
SOURCE bread wheat.
ORGANISM Mitochondrion Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1159)
AUTHORS Breiman,A.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1994) A. Breiman, Tel Aviv University, Tel Aviv,
69978, ISRAEL
REFERENCE 2 (bases 1 to 1159)
AUTHORS Elkeles,A., Devos,K.M., Graur,D., Zizi,M. and Breiman,A.
```

Multiple clones for wheat voltage-dependent anion channels (VDAC): isolation, differential expression, mapping and evolution	
Journal	Plant Mol. Biol. 29 (1), 109-124 (1995)
Medline	96017618
FEATURES	
source	Location/Qualifiers
5'UTR	1..1159
gene	/organism="Triticum aestivum"
	/organelle="mitochondrion"
	/cultivar="Atir"
	/db_xref="taxon:4565"
	/tissue_type="roots"
	/clone_lib="lambda gt11"
	1..51
	52..879
	/gene="VDAC 1"
CDS	52..879
	/gene="VDAC 1"
	/codon_start=1
	/product="voltage dependent anion channel (VDAC)"
	/protein_id="CAA54788.1"
	/db_xref="GI:456672"
	/db_xref="SWISS-PROT:P46274"
	/translation="MGPGGLYSGIGKAKDLYRDYQTDHKFTLTYYTANGRAITATS TKKADLVEIQSIQINKNITYDVVKANSANVITTTIADDLAAPGLKTLISFAVPOQR SGRVEIQYLHDYINASIGLTANPVVNLGSAEGTSALAVGADVSLDTATPKNAKYN AUSYTNODLIASLNKNGDSLTYHYHVEKSGTAVGAELTHFSNNSNLSLTFGQTH TDLPLIVKARINNSGKASALIQHEFMPKSLCTISAEDVTKAIEKSKVGIATALKP"
3'UTR	877..1159
BASE COUNT	308 a 321 c 270 g 260 t
ORIGIN	
alignment_scores:	
Quality:	72.00 Length: 121
Ratio:	1.358 Gaps: 4
Percent Similarity:	43.802 Percent Identity: 24.793

**TITLE**  
JOURNAL  
Direct Submission  
Submitted (31-JAN-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
**FEATURES**  
source  
1. .i0328  
/organism="Caulobacter crescentus"



```

/db_xref="taxon:69394"
complement(90..2633)
/gene="CC3461"
CDS
complement(90..2633)
/gene="CC3461"
/notes="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="TonB-dependent receptor"
/protein_id="AAK25423.1"
/db_xref="GI:13425181"
/translation="MITLVSNSNALILMPSFRVDFHQDEGPCRCVSGSGNSKNAKNTREEMTVSSGRARLFGASLLGAGLVALPSVATAQADPNIAEIVATKRDATIQDIPFISNAQTERDIQSGAVTLEDLSRNAGLTIQNLPGQSQSVRGVSGAGQVVRQPGKQGVGVLDSEVLSLFTPDVLDLNRVETLRGPQGLFGSGSVGGTIRYITNPKLGVSECTFEANALVDGAFGHVKGAVNIPISDKIAMRAVGYLTYGGFIIDARKEGVDKNDGRTGRGRLSFTIEPNEQNFTRVYVYQIRAGGFNRQETNLFANNTTRPKIQIGEREQYLLDLERFSNDTLADMTANKFDAATLTSVTSISRDITVSDSALTSQSVSDLGYPAAAVLLPSNLRDITDLEQFTOELRLASTDGLQWVIGFYSKDRVYNORLPTTYDYTDYDAVILGAGTSAAVANGFGKDSPNAYLPYDIKOKALFGRVTLNAAYTDIKLQTLTDAGSCSSRVVFNPKAHTKGVAELTARLANGLDIGLSDTVTNAQAKSGFRGGVNDPLNPLCSAQDRAIPGGYQNYDDETLWNVEGVSRRFSGSLRAEFSTVKDGTGAVIGGIREGNRLPSVPKQVSDVTYSKEVRPGVNGYLKASLQHVGNRTFQASDQENNRPLFVSNLAFGATGRVPTAVNLALPSYEIVNLSAGLEMQDGVDLTYVNNLFDENALLSFDRERGRARLGVAVGQPRTMGVTVRKSF"
2758..3945
/gene="CC3462"
2758..3945
/gene="CC3462"
/notes="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="thiolase family protein"
/protein_id="AAK25424.1"
/db_xref="GI:13425182"
/translation="MREAVIVSVARTGLAKSVRGFGFNTHGAAMAGHAIQHAVSRAGLEGAEDVVLGCGPEGATGMVNARNAAMWAGLPVTTSGOTINRFCSGGQAIATAANVYRNDGANVAIGGVESISLVNAGGHMNRPHITEKLMOTHPALWAMIDTADIVAKRYNVSREVDYALRSOORIAAQAAGLKDDEIVPMATKMYVNKEKEESFVYVDKDECNRADTTLLEGASLKPVMGEKGFITAGNASQLSDGAAAVYVNEAKEABKRGTLPLGAFRGVAGCEPDEMGIGVPVFRLLERHGLKVDIDIDIELNEAFASQCLYSRDLRGIDPEKYNVNGSGSTAIGHPFMGTCGARCAGHLLLEGKRRKAKLGVVYVTCIGGGMGAALFEIF"
complement(3998..4279)
/gene="CC3463"
CDS
complement(3998..4279)
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK25425.1"
/db_xref="GI:13425183"
/translation="MMELTRMELVLRQAQDEFLRTEANSRPHPELVEGRLLRGDQHTPSQKPERPETAKAVSGLFQSSTLSDCAVRNOKISNRPAAPLWIA"
complement(4343..7246)
/gene="CC3464"
complement(4343..7246)
/gene="CC3464"
/notes="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="DNA polymerase I"
/protein_id="AAK25426.1"
/db_xref="GI:13425184"
/translation="MLPAMTDAAAPLPNASELTQDGPVAVRLFVLDGSAYLFRAYHALPPLTRKSDGLPVAGVQFCNMLKLLRDMQGDTPTHLAVIWDHSEKTERNALYDQYKARHPPEDLLIQPSPLVREATLAFQVPAIELPGVEADDLIAAYACKARDIGGEALIVSSKDQLMLQGVGSEYDPMKVRIEQVFKFGVYPEKVVVDVQALGDSVDNDPTGAPGIGIKTAQVTEIGDLDTLARAGEIKQPKRRTINFADQIIRLSRALVLDKDDTFLPQPLDALTVREPDKEALAAFLQOMEFRSLARRVGDGSAATPGTLDLPAAPKAPVSVSYMGAAARAAAHVPVKIDHAAYACVRDLATLKAWYAKATKGLVAFDFTETDALSSA

```

```

TAGLCGVSLAIAPGEACYIPISHCEKADGLAFAPADIEQIPLADVIATLKPLLEDPAVLKVAONAKYDIAVLARHGTOVAPIEDTMLISVYVLEAGLHGHGDELSELHGHKPIPKVQAGSEKGOISFKHVALPEATVAYAAEDADVTLRLYHLKPOLARASLSTVYETLERPWAYLAMMENNGVRVDPPEARLLSNFESLRMAQFEARQELVGRPNLSPGQIGDVLFGEQMKGGKKTATQWSTDSVLESLAEHELPRVLLDWRQLKUGKPYTENLIAAIAPSTGRVHTSYLAATTTKRLSSSDPNLQNPVTRTEGRKIRKAFVAAPGHVLISADYSQIELRLAHIGIDIPOLKKAFOGDLIHAMTASEMFDPTIEGMDPIRRRAKAINFGIVYISAFLANOLGISQEGAGVIKTYFERFPGIOAYMDATKAFVREHGVYVTTIFGRKINIPDIEAKSAAHROFAERAAINAPIQGAADVMRMRIMPVALAAGLSTRLMLQVHDELVEAPEAEAEACAVVRVAYMEKAAEPAVALSIPLTVEARAASNWDEAH"
7355..9202
/gene="CC3465"
CDS
7355..9202
/gene="CC3465"
/notes="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="ATP-dependent DNA helicase RecQ"
/protein_id="AAK25427.1"
/db_xref="GI:13425185"
/translation="MYVPPSPELDARDVLRRTFFHADFRGLQAGVIELHTGHSAMAVLPTGGGKSLCYQIPSLIRPGLGLVLSPLIALMQQVQLRQGRVAAERLDSNISMDERSDIMRRIDAGEVDLLYLSPEGLMQPMLDRLSMTPLALIAVDHCVSQWGHDFRPEYRMLGLAEFLDPAPRLAVTATARTRDDIRAEELRQGAEEVDSFARPELALSAREKRGKHDRVVELVLERPGSVIYAGSRDGTKEKLAERLNAEIVPALAYHAGLDKAVRARRLEDLEADAAVMVATIAFGMGVDKPDVYVTHADPPAAIEAYWQEVGRAGRDGQPAEGITLYGSADMAWAARRETREAPDEVKQVSRKLRQFYAMLEGVTCRAAANRYRFGEEGVGHGVCDCVSPPTGIDATQAOKALSVAHRLGRGRVRIEHLMKTKTKDVTPEQAQLPTFGIGREFSOPTRWDLFTLIFEGLLREDPDNDRLIGLGDVEGVQRYRREKVALROTTPADPSGGGRAGGMRKREGRLTIPAEQLLFEALRSRWRKEQAQLQHVPPYVIFHDATLAEITAAARPATLALALGKAGGVGGKLDYRGEAVLKVVREN"
9247..9555
/gene="CC3466"
9247..9555
/gene="CC3466"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK25428.1"
/db_xref="GI:13425186"
/translation="MTHVNSLTPSDITSAYDDIVENTKRAKDAALDAVKASARDTATEVAAQOTRTFAAQARAGAAMTTKQAEAHVHAEDHPMGLGALFALGLGLGFLISRNLAR"
complement(9792..10250)
/gene="CC3467"
CDS
complement(9792..10250)
/gene="CC3467"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK25429.1"
/db_xref="GI:13425187"
/translation="MDVIELAPSRPETTLAVTRGAARLLVDLGYAPLAEVTLPLNGRRADLMALGPKGDVLIVEVKSGLEDFVRDKWGDYAPFCDAFVFAVAFSPFGDILPEDPGLLVADGEGGAVVREAPLTPLAPARRKALTAFGLAARAGVNAERLSL"
BASE COUNT 1665 a 3459 c 3503 g 1701 t
ORIGIN

alignment_scores:
  Quality: 72.00      Length: 141
  Ratio: 1.000      Gaps: 9
Percent Similarity: 51.064      Percent Identity: 25.532

alignment_block:
US-09-528-682-4 x AE006005/rev ..
Align seg 1/1 to reverse of: AE006005      from: 1 to: 10328
3 PhetHrValArgile.....
||||| |||:::

```



2











```

:::   ::|||
1900 TAACATATAAGGC 1913

```

```

seq_name: gb_pl:SPAC4G8

```

```

seq_documentation_block:

```

```

LOCUS      SPAC4G8               36776 bp      DNA      linear      PLN 24-NOV-1999
DEFINITION S.pombe chromosome I cosmid c4G8.
ACCESSION  Z56276
VERSION    Z56276.2 GI:4584696
KEYWORDS   condensation-specific protein homologue; leucyl tRNA synthetase;
            MAK32 homologue; mitochondrial carrier protein; nuclease;
            regulatory protein; SEC61 homologue; serine threonine protein
            kinase; SMP3 homologue; zinc finger protein.
            fission yeast.

```

```

SOURCE

```

```

ORGANISM   Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomycetes.

```

```

REFERENCE

```

```

AUTHORS    Badcock,K., Churcher,C.M., Barrell,B.G., Rajandream,M.A. and
            Walsh,S.V.

```

```

TITLE      Direct Submission

```

```

JOURNAL     Submitted (16-OCT-1995) Schizosaccharomyces pombe chromosome I
            sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
            CB10 1RQ E-mail: barrell@sanger.ac.uk

```

```

COMMENT

```

```

On Apr 14, 1999 this sequence version replaced gi:1022345.
Notes:

```

```

Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.

```

```

(URL, http://www.sanger.ac.uk/Projects/S-pombe)

```

```

Protein coding regions (CDS) have been predicted with the help of
computer analysis using the GeneFinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp5splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.

```

```

CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).

```

```

The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.

```

```

The length in codons is given for each CDS.

```

```

IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.

```

```

Cosmid c4G8 overlapped by c22G7, EMBL:SPAC22G7, Z54328, at the 5',
end and c16C9, EMBL:SPAC16C9, Z54366, at the 3' end.

```

```

FEATURES
source

```

```

1..36776
/organism="Schizosaccharomyces pombe"
/strain="972h-"
/db_xref="taxon:4896"
/chromosome="I"
/map="IL"
/clone="cosmid c4G8"
complement(1..256)
/gene="SPAC4G8.01c"
/feature="SPAC22G7.11c"
complement(1..256)
/partial
/gene="SPAC4G8.01c"
/feature="SPAC4G8.01c"

```

```

gene

```

```

crassa, CON6-NEUCR, condensation specific protein 6, (93
aa), E(): 0.0051, 45.5% identity in 55 aa overlap"
/codon_start=1
/label="SPAC4G8.01c"
/product="hypothetical protein"
/protein_id="CAA91204.1"
/db_xref="GI:1022346"
/db_xref="SWISS-PROT:Q09802"

```

```

CDS

```

```

/translation="MPDPNRVLGKATLHNPVNSQAKERAEDYIESHSSQETGDY
SAQAGGRDLDYEDLCGYDEDAFDONEGLNLVGDSEGFVDD"
1..113
/feature="nominal overlap with SPAC22G7, EM:254328 S. pombe
chromosome 1"
complement(join(1352..1364,1533..1635,1743..1839))
/gene="SPAC4G8.02c"
complement(join(1352..1364,1533..1635,1743..1839))
/feature="SPAC4G8.02c"
/feature="SPAC4G8.02c, len:70, SIMILARITY:Canis familiaris,
S61G-CANFA, protein transport protein sec61 gamma
subunit., (68 aa), fasta scores: opt: 245, E():5.6e-13,
(49.3% identity in 67 aa)"
/codon_start=1
/label="SPAC4G8.02c"
/product="putative protein transport protein sec61-gamma
subunit"
/protein_id="CAA91203.1"
/db_xref="GI:1022347"
/db_xref="SWISS-PROT:Q09827"
/translation="MADNADDLFOIPKNEYKSGSHFIKRCVKPDKRFLSISKAVATG
FVLMGLIGYIIKLHPIPKVLVGA"
complement(1365..1382)
/gene="SPAC4G8.02c"
/feature="splice branch and acceptor sequence,
tactaataataat tag"
complement(1365..1532)
/feature="SPAC4G8.02c"
/feature="confirmed intron"
complement(1527..1532)
/feature="SPAC4G8.02c"
/feature="splice donor sequence, gtaagt"
complement(join(1532..1635,1743..1800))
/feature="SPAC4G8.02c"
/feature="Match to PF00584 SecE, SecE/Sec61-gamma subunits of
protein translocation complex Score 57.34"
complement(1636..1649)
/feature="SPAC4G8.02c"
/feature="splice branch and acceptor sequence,
tactaataagaacag"
complement(1636..1742)
/feature="SPAC4G8.02c"
/feature="confirmed intron"
complement(1737..1742)
/feature="SPAC4G8.02c"
/feature="splice donor sequence, gtagcg"
complement(2962..5304)
/feature="SPAC4G8.03c"
complement(2962..5304)
/feature="SPAC4G8.03c"
/feature="SPAC4G8.03c, len:780,
SIMILARITY:Schizosaccharomyces pombe, YDHE_SCHPO,
hypothetical 73.3 kd protein c6g9.14 in chromosome 1.,
(681 aa), fasta scores: opt: 858, E():0, (32.3% identity
in 598 aa)"
/codon_start=1
/label="SPAC4G8.03c"
/product="pumilio domain containing protein"
/protein_id="CAA91204.1"
/db_xref="GI:1022348"
/db_xref="SWISS-PROT:Q09829"
/translation="MVNRDAYNELNLNKKSOETNRKPSPLSYTSISRELDYANOSPF
SSNSFSPTELKARTSPTDVRASSTPLNADPHKYSHDPQRLFKAGNASFSTKS
FPSSVTVSPSEETPLTSGMKNKSVEHYFTLSESAISSHKSSIPERNFSDSSVSVN
PLLHWNVNITLLRQDLENVANSRQDLPYKTFSTTSNSDVLHRESFSFLDDEES
KLASLRNINDRPLPYLKNSERLHRLQLLNHPFSQNNVSTNSKNYSTDTFK
IQSSLNLQNRQNHRIETDQLSHFDHLDPSRIPSPYQPSLQPLESKLHKSVDVH
SKLNLSQLNPILRSENLQNDHNSDMNDPTNVSTKNRNOTVGEHPYVDNK
KKKGPAKPEKATLGRVNSFFGSHSTSNYSKVPLSAKLTGEKSDLSNLLKNGKK
KSDNQIPHLVFLGHLGCKRYLQKLLDENPKVNASLFFPIRQSVQVQMI
DPFGNMQCKLFFVASRQKLSMLNGIGEVIDICSNLYGTRSMQNTIDKTSVQVQIS
LLLKIITPLTTLACDNNGTHVLQKCIAPFKPEKLEPLFLSNEENLITLATNRHGCCII
LQRLDRTNGDIQERLVNSIIKSCULLVQNAVGNVLYQHVLELNLIQPYTERIEKFFG

```

```

misc_feature

```

```

gene

```

```

CDS

```

```

misc_feature

```

```

intron

```

```

misc_feature

```

```

misc_feature

```

```

misc_feature

```

```

intron

```

```

misc_feature

```

```

gene

```

```

CDS

```















OM of: US-09-528-682-4 to: N\_Geneseq\_032802: \* out\_format : pfs

Date: Jun 18, 2002 7:54 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09528682/runat\_18062002\_082444\_7843/app\_query.fasta\_1.689  
-DB=N\_Geneseq\_032802 -QPMF=fastap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.500 -MINMATCH=0.100 -LOPOCL=0.000 -LOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -LOPOCL=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682.CCNI\_1\_677 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-4

Query length: 110

Database: N\_Geneseq\_032802: \*

Database sequences: 1736436

Database length: 85845721

Search time (sec): 554.690000

score\_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF28544 +	6	132.09	52.12	81.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF28528 +	7	133.38	44.13	78.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAF12246 +	15	161.60	1.18	77.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA555859 +	2	155.93	2.45	77.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA60446 +	26	153.46	3.36	76.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA554875 +	65	143.82	11.57	74.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA18499 +	12	157.73	1.94	74.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA18500 +	33	146.73	7.97	74.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ57972 +	74	146.11	8.63	73.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA552186 +	61	161.20	1.25	73.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA552186 +	32	143.89	11.47	73.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA552185 +	32	143.83	11.56	73.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA558840 +	58	87.09	1.7e+04	72.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ90510 +	7	131.36	57.23	71.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV25812 +	74	131.35	57.24	71.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV12115 +	74	131.35	57.24	71.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA91164 +	74	131.35	57.24	71.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA171717 +	33	138.70	22.32	71.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV01866 +	33	138.70	22.32	71.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV50079 +	33	138.70	22.32	71.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA552160 +	3	138.27	23.56	70.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA551503 +	3	138.26	23.60	70.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1988.DAT:AA82053 +	32	137.53	25.93	70.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74365 +	13	122.54	177.24	70.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA91726 +	26	138.72	22.26	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA91599 +	26	138.62	22.53	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA229213 -	26	138.61	22.58	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAD17442 -	3	136.51	29.55	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA229207 -	35	135.42	33.96	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA229219 -	40	133.98	40.85	70.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAQ90932 -	66	151.78	4.17	69.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63571 -	9	148.54	6.32	69.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAQ90931 -	92	148.27	6.54	69.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF27588 +	1	145.87	8.90	69.50	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH41227 -	2	89.11	1.3e+04	69.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH14036 +	1	143.11	12.67	69.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH34611 +	19	139.30	20.66	69.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAQ80898 +	24	137.64	25.57	69.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC81710 +	33	133.75	42.10	69.00	
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA10739 +	28	133.96	40.98	68.50	

/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL25579 +	68.50	129.78	70.04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL25578 -	68.50	123.35	159.83
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAF13281 -	68.00	143.27	12.42
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH72551 +	68.00	137.09	27.42
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA569724 +	68.00	128.55	82.06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAT0204 +	68.00	128.48	82.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24044 +	68.00	126.38	108.32
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH88704 +	68.00	90.64	1.1e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH52867 +	67.50	147.23	7.47
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH82337 +	67.50	140.53	17.63
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ13334 +	67.50	137.35	26.53
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA534420 +	67.50	131.81	53.98
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA534421 +	67.50	131.81	53.98
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA54753 +	67.50	130.49	63.98
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC1665 +	67.00	139.61	19.86
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ4133 +	67.00	136.79	28.52
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC1739 +	67.00	135.28	34.70
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAC13988 +	67.00	126.88	101.61
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA520252 -	67.00	103.97	1.9e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ3838 +	66.50	136.85	28.28
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA552248 +	66.50	128.28	84.86
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA555348 +	66.50	128.09	86.97
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA13289 +	66.50	126.59	105.47
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA552002 +	66.50	120.89	219.01
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1989.DAT:AAAN1695 +	66.50	114.78	479.66
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA554875 +	66.50	110.20	852.87
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA555261 +	66.50	109.36	961.31
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74370 +	66.50	104.12	1.9e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ4165 +	66.00	143.68	11.79
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC36434 +	66.00	139.09	21.24
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL07353 +	66.00	137.15	27.23
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV29572 +	66.00	130.56	63.38
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL08335 +	66.00	127.66	91.93
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL08473 +	66.00	127.39	95.21
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA52198 +	66.00	126.75	103.28
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ49507 +	66.00	122.14	186.67
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA520197 +	66.00	122.13	186.83
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ49509 +	66.00	115.49	438.05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA520197 +	66.00	115.49	438.05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL15995 +	66.00	113.27	582.40
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL15994 -	66.00	100.24	3.1e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ46148 +	65.50	137.37	26.46
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV93875 +	65.50	130.39	64.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV65118 +	65.50	126.45	107.31
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA539169 -	65.50	114.75	481.77
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL17475 +	65.00	139.84	19.28
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA568038 +	65.00	131.27	57.88
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA530596 +	65.00	127.02	99.81
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA575781 +	65.00	127.02	99.81
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA530595 +	65.00	127.01	99.97
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA575780 +	65.00	127.01	99.97
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV42981 +	65.00	122.82	171.04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH73132 +	64.50	138.06	24.21
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA53209 +	64.50	137.23	26.95
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AACT6543 +	64.50	133.23	45.00
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV42337 +	64.50	124.26	142.28
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV64558 +	64.50	124.17	143.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV44449 +	64.50	124.17	143.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA5219359 +	64.50	124.17	143.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA219147 +	64.50	124.17	143.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA54316 +	64.50	123.60	154.82
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH5491 +	64.50	123.37	159.39
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH5491 +	64.50	118.57	295.14
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL02619 +	64.50	112.93	607.98
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL02618 -	64.50	107.26	1.3e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA520251 +	64.50	93.49	7.4e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA161371 -	64.50	74.29	8.5e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA161372 -	64.50	74.29	8.5e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV36099 -	64.00	139.62	19.83
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA568411 -	64.00	132.56	49.04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV25134 +	64.00	125.43	122.32



/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1997.DAT:AAV24920 + 64.00 125.43 122.32 22  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:ABL14596 + 64.00 125.06 128.34 23  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:ABL16369 + 64.00 123.88 149.31 23  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAAS89526 + 64.00 121.07 214.15 33  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAAS89114 + 64.00 121.06 214.23 33  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAV06946 + 64.00 114.79 479.07 62  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAF59105 + 64.00 113.63 555.61 75  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAZ53859 + 64.00 112.76 621.31 75  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAAF1481 + 64.00 92.62 4.8e+03 34  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAAF21612 + 64.00 76.76 1.1e+05 34  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAAF1489 + 64.00 63.50 3.3e+05 83  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAAF1823 + 63.50 138.19 2.1e+05 83  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAAF12239 + 63.50 132.19 51.44 10  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1995.DAT:AAH52360 + 63.50 131.55 55.83 10  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2002.DAT:AAH39124 + 63.50 129.96 68.48 12  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH33159 + 63.50 126.74 103.41 17  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1989.DAT:AAH90684 + 63.50 125.82 116.37 17  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1993.DAT:AAQ47816 + 63.50 125.71 118.07 19  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAH33391 + 63.50 122.58 176.46 19  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH54038 + 63.50 121.54 201.61 22  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:ABL20616 + 63.50 108.42 1.1e+03 29  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2002.DAT:ABAO3041 + 63.50 49.11 1.6e+06 29  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH58859 + 63.00 125.74 117.57 17  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH43327 + 63.00 125.42 122.54 17  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH40308 + 63.00 124.68 134.72 19  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH43236 + 63.00 124.34 140.72 19  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH411430 + 63.00 123.56 155.58 22  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH43235 + 63.00 123.28 161.21 26  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH52192 + 63.00 120.01 245.10 26  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH52191 + 63.00 119.95 247.09 29  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAH88835 + 63.00 118.50 297.76 35  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1996.DAT:AAH40235 + 63.00 118.19 309.65 35  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAV40370 + 63.00 117.62 333.07 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAV40370 + 63.00 117.62 333.07 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH6180 + 63.00 117.62 333.07 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH57049 + 63.00 117.62 333.07 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:ABL15715 + 63.00 117.55 336.17 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH33132 + 63.00 116.99 361.29 37  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAV33774 + 63.00 115.30 448.68 46  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAH33426 + 63.00 113.84 541.05 57  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2002.DAT:AAH33426 + 63.00 113.22 585.53 57  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59101 + 63.00 112.73 624.16 57  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59103 + 63.00 112.44 647.38 60  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59106 + 63.00 112.25 663.69 60  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1996.DAT:AAH38740 + 63.00 111.12 766.91 69  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59102 + 63.00 111.12 767.05 69  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59100 + 63.00 111.12 767.20 69  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH59100 + 63.00 107.84 1.2e+03 69  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAZ45680 + 63.00 107.01 1.3e+03 10  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:ABL15714 + 63.00 106.98 1.3e+03 11  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH10263 + 63.00 105.54 1.6e+03 11  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1990.DAT:AAH404668 + 63.00 105.41 1.6e+03 12  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH28545 + 63.00 88.11 1.5e+04 6  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH28542 + 63.00 87.45 1.5e+04 6  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAH14226 + 63.00 70.15 1.4e+05 3  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT:AAH70203 + 62.50 145.94 8.81 40  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH68783 + 62.50 145.17 9.73 40  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH29112 + 62.50 139.43 20.33 44  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAZ99654 + 62.50 138.62 22.54 44  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH7891 + 62.50 137.31 26.68 50  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH28629 + 62.50 137.31 26.68 50  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH09169 + 62.50 133.38 44.16 73  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH65928 + 62.50 133.34 44.40 73  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH20013 + 62.50 131.68 54.90 86  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH20012 + 62.50 130.63 62.84 96  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH91406 + 62.50 130.20 66.38 96  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT:AAH29530 + 62.50 129.41 73.42 96  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1997.DAT:AAH59539 + 62.50 129.08 76.64 111  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1997.DAT:AAH54888 + 62.50 129.00 77.38 111  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAH16676 + 62.50 129.00 77.38 111  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT:AAH29785 + 62.50 128.85 78.95 111  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT:AAH27553 + 62.50 128.44 83.17 111  
/SIDS5/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:AAH43498 + 62.50 126.72 103.66 111



/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68398 +	61.00	120.07	243.33	60.00	127.46	94.29
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ADA45472 +	61.00	119.53	260.98	60.00	123.57	155.44
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA5568 +	61.00	119.53	260.98	60.00	123.57	155.84
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI14244 +	61.00	119.53	260.98	60.00	122.44	175.64
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI05626 +	61.00	119.53	260.98	60.00	122.44	179.64
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI04081 +	61.00	119.53	260.98	60.00	121.07	214.17
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH89730 +	61.00	119.35	266.93	60.00	120.79	222.00
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA50598 +	61.00	117.60	334.00	60.00	120.66	225.51
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA68554 +	61.00	117.60	334.00	60.00	120.23	238.55
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI23448 +	61.00	117.60	334.00	60.00	118.01	316.89
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI48769 +	61.00	117.60	334.00	60.00	115.87	416.14
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV209075 +	61.00	117.42	342.07	60.00	115.56	434.17
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV25038 +	61.00	117.42	342.07	60.00	113.74	547.77
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV57017 +	61.00	117.35	344.96	60.00	113.57	559.91
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV25013 +	61.00	117.35	344.96	60.00	111.05	773.90
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV57088 +	61.00	117.35	344.96	60.00	108.73	1.0e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAI14100 +	61.00	115.39	443.36	60.00	108.73	1.0e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAO71484 +	61.00	113.81	542.92	60.00	108.73	1.0e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO9668 +	61.00	113.23	584.93	60.00	108.73	1.0e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL22970 +	61.00	112.48	644.20	60.00	108.15	1.1e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAI14547 +	61.00	103.81	2.0e+03	60.00	107.39	1.2e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52031 +	61.00	103.81	2.0e+03	60.00	107.30	1.3e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2655 +	61.00	93.45	7.4e+03	60.00	107.06	1.3e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO2654 +	61.00	86.90	1.7e+04	60.00	106.95	1.3e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH64966 +	61.00	65.21	2.7e+05	60.00	106.72	1.3e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68528 +	61.00	65.21	2.7e+05	60.00	99.99	3.2e+03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH74684 +	60.50	128.52	82.29	60.00	90.45	1.1e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH90731 +	60.50	128.52	82.29	60.00	87.52	1.6e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH50525 +	60.50	128.52	82.29	60.00	86.80	1.7e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH50846 +	60.50	128.52	82.29	60.00	76.27	6.6e+04
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH90844 +	60.50	128.19	85.93	60.00	73.07	1.0e+05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH84131 +	60.50	125.14	127.08	60.00	65.33	2.6e+05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH92542 +	60.50	124.57	136.61	60.00	49.60	1.5e+06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH54646 +	60.50	124.29	141.61	60.00	40.46	3.0e+06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH54782 +	60.50	124.29	141.61	60.00	36.25	3.3e+06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABAO4296 +	60.50	124.08	145.60	60.00	36.24	3.3e+06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH98542 +	60.50	123.39	159.01	60.00	36.24	3.3e+06
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH30801 +	60.50	123.36	159.55	60.00	127.43	94.72
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH37406 +	60.50	123.28	161.33	60.00	126.76	103.13
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH54701 +	60.50	122.96	167.99	60.00	126.58	105.56
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH160712 +	60.50	122.93	168.72	60.00	125.46	121.94
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH58926 +	60.50	122.38	180.96	60.00	125.22	135.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH59133 +	60.50	121.89	192.67	60.00	121.20	210.46
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH1481 +	60.50	121.63	199.25	60.00	120.88	219.24
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH65938 +	60.50	121.40	205.13	60.00	120.70	224.40
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH15353 +	60.50	120.41	233.00	60.00	120.49	230.59
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH42513 +	60.50	119.50	261.74	60.00	120.49	230.59
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH42513 +	60.50	117.32	346.23	60.00	120.47	231.08
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH93446 +	60.50	115.92	414.59	60.00	120.20	239.32
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH93446 +	60.50	115.12	459.35	60.00	120.10	242.33
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH12972 +	60.50	114.75	481.69	60.00	120.05	244.09
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH08330 +	60.50	114.73	482.64	60.00	119.02	278.43
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABH199771 +	60.50	114.61	490.46	60.00	118.42	300.80
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH65243 +	60.50	111.92	692.29	60.00	117.80	325.76
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLO3849 +	60.50	111.64	717.57	60.00	116.41	389.03
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH541070 +	60.50	109.93	893.28	60.00	115.38	444.16
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH20273 +	60.50	109.15	986.95	60.00	115.16	457.11
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABH16561 +	60.50	108.32	1.1e+03	60.00	114.60	490.82
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH51671 +	60.50	107.41	1.2e+03	60.00	114.18	517.79
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABH08331 +	60.50	107.37	1.2e+03	60.00	114.06	526.05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH80200 +	60.50	105.16	1.6e+03	60.00	114.06	526.05
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH99922 +	60.50	105.14	1.7e+03	60.00	113.36	575.63
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH10307 +	60.50	105.10	1.7e+03	60.00	113.25	583.80
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH120119 +	60.50	104.62	1.8e+03	60.00	113.12	593.26
/SID55/cgcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAH52140 +	60.50	100.74	2.9e+03	60.00	111.86	697.46
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68526 +	60.50	94.08	6.8e+03	60.00	111.86	697.78
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH11609 +	60.00	132.17	51.54	60.00	111.07	771.82
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH62399 +	60.00	130.38	64.90	60.00	111.07	771.82
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH68407 +	60.00	129.27	74.82	60.00	110.25	857.42
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH91000 +	60.00	129.27	74.82	60.00	109.59	933.82
/SID55/cgcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH12481 +	60.00	128.63	81.19	60.00	109.59	933.89











/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA60557 + 58.00 106.52 1.4e+03 33  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA96307 + 58.00 106.02 1.5e+03 34  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA541616 + 58.00 105.71 1.5e+03 35  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ5288 + 58.00 105.05 1.7e+03 36  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH54467 + 58.00 104.77 1.7e+03 37  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAH13950 + 58.00 104.39 1.8e+03 38  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH13978 + 58.00 104.39 1.8e+03 39  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV16171 + 58.00 104.39 1.8e+03 40  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH58898 + 58.00 104.12 1.9e+03 41  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH57486 + 58.00 104.06 1.9e+03 42  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH54213 + 58.00 103.97 1.9e+03 43  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH09904 - 58.00 102.67 2.3e+03 44  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH09078 + 58.00 101.97 2.5e+03 45  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH3272 + 58.00 101.83 2.5e+03 46  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH12178 - 58.00 101.54 2.6e+03 47  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH70184 + 58.00 101.19 2.7e+03 48  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABH06180 + 58.00 100.60 3.0e+03 49  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH23096 + 58.00 100.04 3.2e+03 50  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH16441 + 58.00 98.29 4.0e+03 51  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA09696 + 58.00 97.80 4.2e+03 52  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH11678 + 58.00 96.37 5.1e+03 53  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH74989 + 58.00 96.31 5.1e+03 54  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH11677 + 58.00 96.25 5.2e+03 55  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH76401 + 58.00 95.96 5.4e+03 56  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABA06648 + 58.00 95.65 5.6e+03 57  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH90837 + 58.00 94.79 6.2e+03 58  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH05156 - 58.00 94.52 6.4e+03 59  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH19201 + 58.00 92.13 8.8e+03 60  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH19680 - 58.00 90.59 1.1e+04 61  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:ABH32624 - 58.00 89.94 1.2e+04 62  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH19200 + 58.00 89.65 1.2e+04 63  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABH15437 + 58.00 87.19 1.7e+04 64  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH75998 + 58.00 86.23 2.9e+04 65  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH829212 + 58.00 82.71 5.2e+04 66  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH81458 + 58.00 78.21 5.2e+04 67  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH21613 + 58.00 65.82 2.7e+05 68  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH11224 - 57.50 124.59 136.24 69  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH87662 + 57.50 122.52 177.71 70  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH12460 + 57.50 122.49 178.45 71  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ04968 + 57.50 118.48 298.50 72  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ72260 + 57.50 118.48 298.50 73  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ01980 + 57.50 115.95 413.05 74  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH21221 + 57.50 115.34 446.16 75  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAH74745 - 57.50 115.24 452.30 76  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH21219 + 57.50 114.67 486.65 77  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH86747 + 57.50 114.46 499.69 78  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH18253 + 57.50 113.40 572.22 79  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH31558 + 57.50 113.40 572.22 80  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH86594 + 57.50 112.77 620.39 81  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH72749 + 57.50 112.77 620.39 82  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH72775 + 57.50 112.77 634.78 83  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABH14617 + 57.50 112.57 636.84 84  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH82305 - 57.50 112.57 636.84 85  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH82306 - 57.50 112.57 636.84 86  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH12090 + 57.50 101.61 818.61 87  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:AAH90389 + 57.50 109.44 951.46 88  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH27638 + 57.50 108.20 1.1e+03 89  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH86435 + 57.50 107.06 1.3e+03 90  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH86411 + 57.50 106.92 1.4e+03 91  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH91696 + 57.50 106.21 1.4e+03 92  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH91695 + 57.50 106.18 1.4e+03 93  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH91571 + 57.50 105.89 1.5e+03 94  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH54334 + 57.50 105.85 1.5e+03 95  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH52184 + 57.50 105.40 1.6e+03 96  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:AAH91089 + 57.50 105.25 1.6e+03 97  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAH23700 + 57.50 103.69 2.0e+03 98  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAH82995 + 57.50 103.69 2.0e+03 99  
/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAH04132 + 57.50 103.24 2.1e+03 100  
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABH21222 - 57.50 102.87 2.2e+03 101



/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52138-	57.00	89.07	1.3e+04	13	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF59104+	56.50	96.37	5.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52168-	57.00	85.18	2.0e+04	19	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABD10021+	56.50	96.34	5.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV52168-	57.00	81.68	3.0e+04	29	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABD11002+	56.50	95.36	5.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV28526-	57.00	80.38	3.9e+04	33	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV70114+	56.50	95.20	5.9e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV5128+	57.00	79.74	4.3e+04	33	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV74783+	56.50	92.92	7.9e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV20249+	57.00	66.86	2.2e+05	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABL13600+	56.50	92.22	8.7e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV20249+	57.00	66.86	2.2e+05	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74427+	56.50	91.11	1.0e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV26869-	57.00	56.77	7.3e+05	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV79779+	56.50	90.97	1.0e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV21610+	57.00	55.34	8.5e+05	34	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV06749+	56.50	87.76	1.5e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV20248+	57.00	45.33	2.2e+06	91	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ56386+	56.50	87.76	1.5e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV20248+	57.00	40.56	2.9e+06	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAZ10086+	56.50	87.42	1.6e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV39331+	56.50	123.92	148.44	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV33304+	56.50	87.37	1.6e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV57081+	56.50	123.62	154.33	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV09930-	56.50	85.10	2.2e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV36622+	56.50	123.62	154.33	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV13047+	56.50	83.24	2.7e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24755+	56.50	122.06	188.43	52	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV13060+	56.50	78.88	4.8e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV23279-	56.50	119.43	264.07	67	/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ23902-	56.50	74.46	8.3e+04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV568719+	56.50	119.36	266.68	6	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH02340+	56.50	62.10	3.8e+05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV120101+	56.50	118.52	296.32	7	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH02339+	56.50	62.16	3.9e+05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV09444+	56.50	117.82	324.65	79	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH68528-	56.50	54.10	9.8e+05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV33193+	56.50	116.31	394.04	8	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV89719+	56.00	123.95	148.04
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV33457+	56.50	115.91	414.98	95	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ25628-	56.00	121.74	196.44
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH66471-	56.50	115.84	418.50	9	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV97720-	56.00	121.25	209.06
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV05751+	56.50	115.04	463.68	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV51463+	56.00	120.37	234.33
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV47375+	56.50	114.94	469.72	10	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ18023+	56.00	120.25	237.75
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV85191+	56.50	114.94	469.72	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ07882-	56.00	119.58	259.14
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV36841+	56.50	114.91	471.53	10	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV34035+	56.00	119.34	267.36
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV72281+	56.50	114.78	479.41	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV10205+	56.00	119.32	267.95
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV31304+	56.50	113.88	538.60	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV36013+	56.00	116.91	364.97
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV32273+	56.50	113.87	539.23	11	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH31759-	56.00	116.20	399.81
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV32365+	56.50	112.72	624.72	12	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV32679+	56.00	115.50	437.44
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV47803+	56.50	112.62	632.53	13	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH05267-	56.00	115.49	438.11
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV46258+	56.50	111.77	705.91	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV34596-	56.00	115.20	454.22
/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAV92525+	56.50	111.64	717.33	14	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ12552-	56.00	114.81	477.97
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV20702+	56.50	111.08	771.00	15	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ04839+	56.00	113.62	556.69
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV39565+	56.50	110.96	782.67	15	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV31679+	56.00	113.62	556.69
/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAV3136+	56.50	110.76	803.39	15	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV31623+	56.00	113.58	559.53
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV86664+	56.50	110.01	884.75	16	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAJ1209+	56.00	113.37	574.51
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV586042+	56.50	108.82	1.0e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV163870+	56.00	113.37	574.51
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV424013+	56.50	108.79	1.0e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV163870+	56.00	113.37	574.51
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV17209+	56.50	108.79	1.0e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV31636+	56.00	113.37	574.51
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU21007+	56.50	108.79	1.0e+03	1	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV31636+	56.00	112.60	634.69
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV30799+	56.50	108.64	1.1e+03	19	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV57670+	56.00	112.53	639.83
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV49168+	56.50	108.44	1.1e+03	19	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV355564+	56.00	110.67	812.15
/SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAV44510+	56.50	107.48	1.2e+03	22	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV39114+	56.00	110.67	812.15
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV28606+	56.50	106.88	1.3e+03	22	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV32973+	56.00	110.55	824.68
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU09455+	56.50	106.88	1.3e+03	22	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV49703+	56.00	110.45	835.88
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU09455+	56.50	106.83	1.3e+03	24	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV40085-	56.00	110.42	838.83
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV63223+	56.50	106.23	1.4e+03	24	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV45741+	56.00	110.29	853.05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV165790+	56.50	106.23	1.4e+03	24	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV885402+	56.00	109.20	981.21
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV588411-	56.50	105.81	1.5e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV14022-	56.00	108.99	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV588411-	56.50	105.81	1.5e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV00699+	56.00	108.89	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV34687+	56.50	105.78	1.5e+03	25	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV1712+	56.00	108.89	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV74416-	56.50	105.16	1.6e+03	26	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV1712+	56.00	108.85	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV20101B.DAT:AAV227746+	56.50	104.84	1.7e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAQ13407-	56.00	108.68	1.0e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV20101B.DAT:AAV227746+	56.50	104.67	1.8e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV33677+	56.00	108.56	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV10022+	56.50	104.67	1.8e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV33677+	56.00	108.37	1.1e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV98375+	56.50	104.53	1.8e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV30269+	56.00	107.61	1.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV51819+	56.50	104.11	1.9e+03	2	/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ29189+	56.00	107.53	1.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV51819+	56.50	103.87	1.9e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ80910+	56.00	107.53	1.2e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV16621+	56.50	103.59	2.0e+03	30	/SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ80913+	56.00	107.24	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV16621+	56.50	103.56	2.0e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAV62577-	56.00	107.23	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV24176+	56.50	102.37	2.4e+03	3	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV51620+	56.00	107.16	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV52176+	56.50	101.80	2.5e+03	36	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV25635+	56.00	106.99	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV32175+	56.50	101.75	2.6e+03	36	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV25635+	56.00	106.92	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV24936+	56.50	101.24	2.7e+03	38	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV30000+	56.00	106.72	1.3e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV25094+	56.50	101.24	2.7e+03	38	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV30892+	56.00	106.71	1.4e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV77593+	56.50	101.03	2.8e+03	39	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV39222+	56.00	105.95	1.5e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV52261+	56.50	100.76	2.9e+03	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU14955+	56.00	105.93	1.5e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV55402+	56.50	100.75	2.9e+03	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAV25376+	56.00	105.26	1.6e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV11003+	56.50	100.66	2.9e+03	4	/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU15131+	56.00	104.59	1.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV86663+	56.50	99.80	3.3e+03	44	/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAV51621-	56.00	104.59	1.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU16620+	56.50	97.94	4.2e+03	5	/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV67228-	56.00	104.59	1.8e+03
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABU05234-	56.50	96.61	4.9e+03	6				







CC infection and life-threatening, systemic diseases including endocarditis  
XX and meningitis.

SQ Sequence 28626 BP; 8524 A; 5568 C; 6119 G; 8415 T; 0 other;

## alignment\_scores:

Quality: 78.00 Length: 116  
Ratio: 1.300 Gaps: 6  
Percent Similarity: 51.724 Percent Identity: 29.310

## alignment\_block:

US-09-528-682-4 x AAF28528 ..

Align seg 1/1 to: AAF28528 from: 1 to: 28626

5 ValArgIleGlnAspAlaTyrGluProIleAla.....AsnThrAs 18

8394 GTTAGCAGCTCGTGATGCCTAAATATGATTGCTAAATTTTCGTAACCA 8443

18 nThrThrThrGlnPheLeuAsnMetGly.....AsnGluValAlaLeuA 33

8444 CCNAAACAGCCCTGTTTATGATGGTTATTAAACCTGTGGAATCA 8493

33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyMetAsp 49

8494 TTGGT...TATGATAATTTTGTCTAATTTGTGCTGATGCGCGTTGAT 8540

50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG 66

8541 GCAATTTTGTATG.....GTGGAT..... 8558

66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83

8559 .....TTACCCGCCCAACGACACATGGCTATGCAAAATCACTGA 8598

83 snLysGlnHisAspAlaAspPheLysLysTyrIle..... 94

8599 CCGAGCGTACCGATCATCCCATGAATCAAAATTTTGTGCTGGCACCAC 8648

95 .....LysLeuLeuIleAsnAsnAspGlyPhe 103

8649 ACAAAATCAGATCGCGGTACCAAGGTTATCGAAACTGTAGTGGGTTT 8696

seq\_name: /SID85/sgcdata/geneseq/geneseqn-emb1/NA2000.DAT:AAF12246

## seq\_documentation\_block:

ID AAF12246 standard; cDNA; 1523 BP.

AC AAF12246;

DT 13-MAR-2001 (first entry)

DE Aspergillus oryzae EST SEQ ID NO:4769.

XX Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.

OS Aspergillus oryzae.

XX WO200056762-A2.

PN 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07781.

XX 22-MAR-1999; 99US-0273623.

XX (NOVO ) NOVO NORDISK BIOTECH INC.

PA (NOVO ) NOVO NORDISK AS.

XX

PI

XX

DR

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC



89 AspPheLysTyrIleLysLeuLeuLeuAsnAspGlyPhePhe 105  
1002 GGATTCAGCCAGGA.....AGCAACAATGACGCAACACAG 1039  
105 rAsnAsn 107  
1040 CAATAAC 1046

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA555859

seq\_documentation\_block:  
ID AA555859 standard; DNA; 2619 BP.

AA555859;  
13-FEB-2002 (first entry)  
DE Streptococcus pneumoniae DNA for cellular proliferation protein #430.  
KW Antisense; ds; prokaryotic cellular proliferation gene;  
KW antibiotic; antibacterial; drug design.  
OS Streptococcus pneumoniae.

XX WO200170955-A2.  
XX 27-SEP-2001.  
XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.  
XX 23-MAY-2000; 2000US-206848P.  
XX 26-MAY-2000; 2000US-207727P.  
XX 23-OCT-2000; 2000US-242578P.  
XX 27-NOV-2000; 2000US-253625P.  
XX 22-DEC-2000; 2000US-257931P.  
XX 16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU38000.

XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 9496; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen  
XX for homologous nucleic acids which are required for cell proliferation in  
XX a wide variety of organisms. The present sequence encodes an  
XX essential prokaryotic cellular proliferation protein.  
XX Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic  
XX format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2619 BP; 768 A; 566 C; 606 G; 679 T; 0 other;

alignment\_scores:  
Quality: 77.00 Length: 94  
Ratio: 1.481 Gaps: 5  
Percent Similarity: 55.319 Percent Identity: 29.787  
alignment\_block:  
US-09-528-682-4 x AA555859/rev ..  
Align seg 1/1 to reverse of: AA555859 from: 1 to: 2619  
14 IleAlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluVa 30  
2028 GTTGCCACACATGAGTACCACCAAGTTC.....AATAGAGT 1991  
30 lAlaLeuAspGlyArgTyrSerAsnTyr...AlaLeuIleSerAlaGluG 46  
1990 AGTCACCGATAGTCAGACGACGAACTTCCTGCCGTATTTCACCAAG 1941  
46 lYglyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62  
1940 AGGCCATAGCTCCCATTTCTTTAGCAGTCAATATC..... 1903  
63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79  
1902 .....CGTTTCAAC..... 1894  
79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96  
1893 ....TGTCTTAACCTCAAGAGCTTCCAGATTTT.....TTCGTTGACT 1854  
96 euleuIleAsnAsnAspGlyPhePheSerAsn 106  
1853 TGTGTGTTCAATCGCAGCGCAATTTCTTCAGCAGT 1822

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA788094

seq\_documentation\_block:

ID AA788094 standard; DNA; 2619 BP.

XX AC AA788094;

XX DT 28-APR-1998 (first entry)

XX DE Streptococcus pneumoniae alanyl tRNA synthetase gene.

XX KW Alanyl tRNA synthetase; alaS polypeptide; immune response; vaccine;  
XX KW antibacterial; meningitis; bacterial infection; ss.

XX OS Streptococcus pneumoniae.

XX FH Key Location/Qualifiers

XX FT CDS 1..2619

XX FT /\*tag= a

XX FT /product= "alaS"

XX PN WO9739013-A1.

XX PD 23-OCT-1997.

XX PF 18-APR-1997; 97WO-US06635.

XX PR 18-APR-1996; 96GB-0007991.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX PI Lawlor EJ;

XX DR WPI; 1997-526389/48.

XX DR P-PSDB; AAW27658.

XX PT Polynucleotide encoding Streptococcus pneumoniae alanyl tRNA



PT synthetase - useful to diagnose, treat or prevent bacterial  
XX infection, especially meningitis

PS Claim 4; Page 37-38; 44pp; English.

XX The present sequence encodes alanyl tRNA synthetase (alas) from  
CC Streptococcus pneumoniae. The polynucleotide encoding alas can be used  
CC for the recombinant production of alas, which can be used to treat  
CC conditions requiring alas activity. Sequences antisense to the  
CC polynucleotide can be used to control alas expression. Alas, or alas  
CC expressing vectors can be used to induce an immune response, i.e. an  
CC antibody (Ab) and/or T cell response, against S. pneumoniae to protect  
CC against infection, or to screen for antagonists or agonists of the  
CC polynucleotide encoding alas or alas's activity, i.e. antibacterials.  
CC The antagonist, e.g. an anti-alas Ab, can be used to treat conditions  
CC requiring alas inhibition, e.g. an S. pneumoniae infection, particularly  
CC meningitis. A fragment of the polynucleotide encoding alas can be used  
CC as a probe to isolate full length or related sequences, or diagnose, e.g.  
CC by polymerase chain reaction, infection stage and type, including  
CC mutation and polymorphism detection. Diagnosis may also be achieved by  
CC detecting alas gene overexpression, e.g. by immunoassay. The Ab can be used  
CC to treat infection, isolate or identify alas expressing clones, purify  
CC alas and as an immunoassay reagent. More generally, the products can  
CC prevent adhesion of bacteria to wounds and in dwelling devices, block  
CC alas protein mediated invasion of mammalian cells and block the normal  
CC progression of infection.

XX SQ Sequence 2619 BP; 766 A; 570 C; 607 G; 676 T; 0 other;

#### alignment\_scores:

Quality: 76.00 Length: 94  
Ratio: 1.490 Gaps: 5  
Percent Similarity: 54.255 Percent Identity: 30.851

#### alignment\_block:

US-09-528-682-4 x AAT88094/rev ..

Align seg 1/1 to reverse of: AAT88094 from: 1 to: 2619

```

14  ILeAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
208  GTTGCACATGATGATACCAACCAAGTTC.....AATAGACT 1991
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
30  lAlaLeuAspGlyArgTyrSerAsn...TyrAlaLeuLeuSerAlaGluG 46
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1990  AGTCACCGATATGTCAGCACAGCACTGCCCTGCCGTATTTCTCACCAGAG 1941
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1940  AGGGCGATAGCTCCCATTTCTTTAGCAGTGTCATATC..... 1903
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
63  GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1902  .....CGTTTCAAC..... 1894
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
79  rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1893  ....TGCTTACTTCAAGAGCTTCCACATTTT.....TTCGTTGACT 1854
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
96  euLeuLeuAsnAsnAspGlyPhePheSerAsn 106
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1853  TGCTGTTCATTCGACGCAATTTCTTCAGCAGT 1822
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

seq\_name: /SIDS5/ycgdata/geneseq/geneseq-emb1/WA2000.DAT.AAA60446

#### seq\_documentation\_block:

ID AAA60446 standard; cDNA; 6585 BP.

XX AAA60446;

XX 09-OCT-2000 (first entry)

XX

DE Murine factor V encoding cDNA SEQ ID NO:4.

XX Murine; factor V; FV; activated protein C; APC; anticoagulant;  
KW activated protein C resistant factor V; thrombosis; screening;  
KW thrombophilia; ds.

OS Mus sp.

XX Key Location/Qualifiers

XX CDS 6..6557

FT /\*tag= a

FT /product= "Factor V"

PN US5066778-A.

PD 23-MAY-2000.

XX 06-NOV-1996; 96US-0746111.

XX 06-NOV-1996; 96US-0746111.

XX (UNMI ) UNIV MICHIGAN.

XX Ginsburg D, Cui J;

XX WPI; 2000-410682/35.

XX P-PSDB; AAB03533.

XX New transgenic mice expressing activated protein C resistant factor V

XX and factor V null transgenic mice useful for screening anticoagulants,  
XX as models for human thrombophilia and as models for testing in utero  
XX gene therapy protocols -

XX Example 1; Fig 2; 76pp; English.

XX The present invention describes transgenic mice (I) and (II) containing  
XX modifications in the factor V gene, where (I) expresses an activated  
XX protein C (APC) resistant factor V and (II) lacks the ability to express  
XX wild-type factor V. The transgenic animals (I) and (II) are useful for  
XX screening compounds with anticoagulant activity. Methods from the present  
XX invention, and the transgenic animals, are also useful in providing  
XX models for human thrombophilia. These models are useful in providing  
XX insight into the basic regulatory mechanisms of blood coagulation and  
XX pathogenesis of human thrombosis. In addition, factor V null transgenic  
XX mice, especially pregnant females may be used as a model system to test  
XX in utero gene replacement therapy protocols. The present sequence  
XX encodes murine factor V, which is used in an example from the present  
XX invention.

XX SQ Sequence 6585 BP; 1946 A; 1675 C; 1432 G; 1532 T; 0 other;

#### alignment\_scores:

Quality: 76.00 Length: 106  
Ratio: 1.169 Gaps: 4  
Percent Similarity: 61.321 Percent Identity: 25.472

#### alignment\_block:

US-09-528-682-4 x AAA60446 ..

Align seg 1/1 to: AAA60446 from: 1 to: 6585

```

1  AspTyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
2058  GATTATGACANTGAGGACTCATATGAGATTATGAACTCTCTGCACCTAC 2107
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
17  rAsnThrThrThrGlnPheLeu.....AsnMetGlyAsnGluValAlaL 32
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
2108  ATCCATGACAACTCGGAGAAATTCATGATTCCTTAGAATAATGATTTGGCA 2157
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
32  euAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMet 48
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
2158  TAGACAACGAAGATGATGATTACCACTACTTACTGCGCTCATCATTTAGGA 2207
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```











840 TGTCTGGTTCGACCTCTCGTGCACCGCCGACCGACAC..... 880  
92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103  
881 .....CTGAGGGCCACCGCCCAACAGCAGCGTTC 910

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAF13012

seq\_documentation\_block:  
ID AAF13012 standard; cDNA; 616 BP.  
XX  
AC AAF13012;  
XX  
XX 13-MAR-2001 (first entry)  
XX  
DE Aspergillus oryzae EST SEQ ID NO:5535.  
XX  
XX Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Aspergillus oryzae.  
XX  
PN WO200056762-A2.  
XX  
XX 28-SEP-2000.  
XX  
XX 22-MAR-2000; 2000WO-US07781.  
XX  
XX 22-MAR-1999; 99US-0273623.  
XX  
XX (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS.  
XX  
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX WPI; 2000-594572/56.  
XX  
XX Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags -  
XX  
XX Claim 88; Page 2299; 3161pp; English.  
XX  
CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring  
CC the global expression of genes from FF cells allows the production  
CC potential of the microorganisms to be improved. New genes may be  
CC discovered, possible functions of unknown open reading frames can be  
CC identified and gene copy number variation and stability can be  
CC monitored. The expression of genes can be used to study how FF cells  
CC adapt to changes in culture conditions, environmental stress, spore  
CC morphogenesis, recombination, metabolic or catabolic pathway  
CC engineering. Using ESTs provides several advantages over genomic or  
CC random cDNA clones including elimination of redundancy as one spot on an  
CC array equals one gene or open reading frame, and organisation of the  
CC microarrays based on function of the gene products to facilitate  
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from  
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus  
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and  
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are  
CC all specifically claimed in the present invention.  
XX  
XX Sequence 616 BP; 172 A; 161 C; 148 G; 135 T; 0 other;

alignment\_scores:  
Quality: 73.00 Length: 136  
Ratio: 1.123 Gaps: 6  
Percent Similarity: 47.794 Percent Identity: 22.794  
alignment\_block:  
US-09-528-682-4 x AAF13012 ..  
Align seg 1/1 to: AAF13012 from: 1 to: 616  
15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31  
:::||||| ||||| ||| |||||::: |||  
110 TCTAATGTGTTTACTACTACACATTCCGAAACGCTGCTCAACGCGATCAA 159  
:::||||| ||||| ||| |||||::: |||  
31 aLeuAspGlyArgTyrSerAsn..... 38  
:::||||| |||||  
160 GATGGAGGCCAAGGCAAGCAATGAGAACTTCAAGCTGGAGAACCTCTTTA 209  
:::||||| ||||| ||| |||||::: |||  
39 .....TyrAlaLeuIleSerAlaGluGly..... 47  
:::||||| ||||| ||| |||||::: |||  
210 ATGTTAAGGCCAAAGTTGCCCTTATTACAGGAGGAGGCTCGGGTATCGGG 259  
:::||||| ||||| ||| |||||::: |||  
47 ..... 47  
260 TTGATGCCACATCAAGCTCTGGCGCTCAACGGAGCCCAAGGCTCTACATCAC 309  
:::||||| ||||| ||| |||||::: |||  
48 .....MetAspArg.....AspLeuPheGlySerA 56  
:::||||| ||||| ||| |||||::: |||  
310 AGGTGCAACTGGCGAGAAGCTCGACCGCGTTGCAGAAATTATAC...AACA 356  
:::||||| ||||| ||| |||||::: |||  
56 laAsnIleAspGlyPheProGluValArgGluPheAsnSerLeuProAsn 72  
:::||||| ||||| ||| |||||::: |||  
357 AGAATCCAGGGC.....GAGATCATCCCCCATC 385  
:::||||| ||||| ||| |||||::: |||  
73 AsnLysAlaSerSerAspThrAlaSerLeuAsnLys.....GlnHisAs 87  
:::||||| ||||| ||| |||||::: |||  
386 ACATCCGACATCACATAAGAGCTCAGTCGACACAGCTGGTGCACAGAGAT 435  
:::||||| ||||| ||| |||||::: |||  
87 pAlaAspPheLysLysTyrIleLysLeuLeuIleAsnAsnAspGlyPheP 104  
:::||||| ||||| ||| |||||::: |||  
436 CTCTCGCGCGAAGTAAGTACCTCTCCATCTCTGATTAACAACGCGCGATCA 485  
:::||||| ||||| ||| |||||::: |||  
104 heSerAsn 106  
:::||||| ||||| ||| |||||::: |||  
486 GCAGTTTCG 493  
seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA52186  
seq\_documentation\_block:  
ID AAA52186 standard; DNA; 3222 BP.  
XX  
AC AAA52186;  
XX  
XX 11-SEP-2000 (first entry)  
XX  
DE DNA encoding H. influenzae strain LDC2 mature hmw2A, SEQ ID NO:48.  
XX  
KW Mature HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;  
KW non-typeable Haemophilus influenzae; NTHi; non-encapsulated;  
KW recombinant production; Escherichia coli; antibacterial; vaccine;  
KW human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;  
KW detection; diagnosis; ds.  
XX  
OS Haemophilus influenzae strain LDC2.  
XX  
PN WO200020609-A2.  
XX  
XX 13-APR-2000.  
XX  
XX 07-OCT-1999; 99WO-CA00938.  
XX



```
PR 07-OCT-1998; 98US-0167568.
PR 08-DEC-1998; 98US-0206942.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Loosmore SM, Yang Y, Klein MH;
XX
DR WPI; 2000-303789/26.
DR P-PSDB; AAB01837.
XX
XX Nucleic acid molecule for producing recombinant high molecular weight
PT proteins of Haemophilus which are used as a vaccine to provide
PT protection against Haemophilus induced diseases in humans -
XX
PS Claim 7; Fig 23A-P; 307pp; English.
XX
CC The invention relates to the recombinant production of Haemophilus
CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The
CC expression construct used to effect recombinant expression comprises a
CC promoter functional in E. coli (e.g., the T7 promoter) operably linked
CC to a modified hmWABC operon from a non-typeable (non-encapsulated) H.
CC influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene
CC clusters termed hmw1ABC and hmw2ABC. Each hmWABC operon comprises hmWA,
CC hmWB and hmWC genes. The hmWA genes encode the structural HMWA proteins
CC and the hmWB and hmWC genes encode accessory proteins which are
CC responsible for post-translational processing and secretion of the HMWA
CC proteins. The modified hmWABC operon used in the expression construct of
CC the invention contains an A gene modified such that it encodes only the
CC mature HMWA. The invention also discloses hmWA genes (AAAS52175-A52198)
CC and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae
CC strains Joysc, KI, K21, LDC2, PMH1, 15 and 12. The nucleic acids and
CC vectors are used for the production of recombinant H. influenzae HMW
CC proteins which can be used as vaccines to mediate a humoral or
CC cell-mediated immune response to provide protection against diseases in
CC humans caused by H. influenzae (e.g., otitis media, epiglottitis,
CC pneumonia and tracheobronchitis). The HMW proteins are also useful as
CC antigens in immunoassays for detecting antibodies against Haemophilus,
CC HMW proteins and/or HMW peptides. The nucleotide sequences encoding the
CC HMW proteins can be used to isolate and clone hmw genes from other
CC non-typeable strains of Haemophilus via hybridisation reactions. The
CC present sequence represents a modified NTHi hmWA gene encoding a mature
CC HMWA protein.
XX
SQ Sequence 3222 BP; 1179 A; 599 C; 659 G; 785 T; 0 other;
```

```
alignment_scores:
  Quality: 73.00      Length: 82
  Ratio: 1.352        Gaps: 3
  Percent Similarity: 65.854  Percent Identity: 28.049
```

alignment\_block:

US-09-528-682-4 x AAAS52186 ..

Align seg 1/1 to: AAAS52186 from: 1 to: 3222

```
28 AsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAl 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
508 AACAAATGTGCTTTAAATGGACGGGTAAAGTCTAAGTATTATTCAAC 557
44 aGluGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyp 61
: : : : : : : : : : : : : : : : : : : : : : : : : :
558 AGCAAGCAATTTATCATAGACTTGACGGTGAATTAATGATCTCGAA 607
61 heProGluValArgGluPhe.....AsnSerLeuProAsnAsnLysAla 75
: : : : : : : : : : : : : : : : : : : : : : : : : :
608 ATGTAACAATTAATCAACACGACGCAAAACATTCGAATCTGGAAGCT 657
76 SerSerAspThr.....AlaSerLeuAsnLysGlnHisAspAl 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
658 AGCAGCGATTCCTTATTGGAAATGTCATCTCTTTAATTTGAGAGAATTC 707
88 aAspPheLys.....LysTyrIleLysLeuLeuIleAsnAsnAsp 101
```

```
seq_name: /SIDSS/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS52185
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
708 AAGAGTTTACCTTTATCAATACGTTAACTCTGCAGAAATGGTGAT 753
```

seq\_documentation\_block:

ID AAA52185 standard; DNA; 3240 BP.

XX AAA52185;

XX DT 11-SEP-2000 (first entry)

XX DE Haemophilus influenzae strain LDC2 hmW2A gene, SEQ ID NO:46.

XX KW HMW protein; hmW gene; hmWA1; hmWA2; high molecular weight;

XX KW non-typeable Haemophilus influenzae; NTHi; non-encapsulated;

XX KW recombinant production; Escherichia coli; antibacterial; vaccine;

XX KW human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;

XX KW detection; diagnosis; ds.

OS Haemophilus influenzae strain LDC2.

PN WO200020609-A2.

XX 13-APR-2000.

XX 07-OCT-1999; 99WO-CA00938.

XX 07-OCT-1999; 98US-0167568.

PR 08-DEC-1999; 98US-0206942.

XX (CONN-) CONNAUGHT LAB LTD.

XX PI Loosmore SM, Yang Y, Klein MH;

XX DR WPI; 2000-303789/26.

XX P-PSDB; AAB01836.

XX Nucleic acid molecule for producing recombinant high molecular weight

PT proteins of Haemophilus which are used as a vaccine to provide

PT protection against Haemophilus induced diseases in humans -

XX Claim 12; Fig 23A-P; 307pp; English.

XX The invention relates to the recombinant production of Haemophilus  
CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The  
CC expression construct used to effect recombinant expression comprises a  
CC promoter functional in E. coli (e.g., the T7 promoter) operably linked  
CC to a modified hmWABC operon from a non-typeable (non-encapsulated) H.  
CC influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene  
CC clusters termed hmw1ABC and hmw2ABC. Each hmWABC operon comprises hmWA,  
CC hmWB and hmWC genes. The hmWA genes encode the structural HMWA proteins  
CC and the hmWB and hmWC genes encode accessory proteins which are  
CC responsible for post-translational processing and secretion of the HMWA  
CC proteins. The modified hmWABC operon used in the expression construct of  
CC the invention contains an A gene modified such that it encodes only the  
CC mature HMWA. The invention also discloses hmWA genes (AAAS52175-A52198)

CC and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae  
CC strains Joysc, KI, K21, LDC2, PMH1, 15 and 12. The nucleic acids and  
CC vectors are used for the production of recombinant H. influenzae HMW  
CC proteins which can be used as vaccines to mediate a humoral or  
CC cell-mediated immune response to provide protection against diseases in  
CC humans caused by H. influenzae (e.g., otitis media, epiglottitis,

CC pneumonia and tracheobronchitis). The HMW proteins are also useful as  
CC antigens in immunoassays for detecting antibodies against Haemophilus,  
CC HMW proteins and/or HMW peptides. The nucleotide sequences encoding the  
CC HMW proteins can be used to isolate and clone hmw genes from other

CC non-typeable strains of Haemophilus via hybridisation reactions. The  
CC present sequence represents an hmWA gene from a non-typeable strain of H.

XX influenzae.

XX Sequence 3240 BP; 1185 A; 600 C; 665 G; 790 T; 0 other;



```
alignment_scores:
  Quality: 73.00      Length: 82
  Ratio: 1.352      Gaps: 3
  Percent Similarity: 65.854      Percent Identity: 28.049

alignment_block:
US-09-528-682-4 x AAA52185 ..
Align seg 1/1 to: AAA52185 from: 1 to: 3240

28 AsnGluValAlaLeuAspGlyArgTySerAsnTyrAlaLeuIleSerAl 44
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
526 AACAAATGCTTTAAATGGACGGGTAATGGTCTAAGTATTATTTCAAC 575

44 aGluGlyClyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyP 61
.: : : : : : : : : : : : : : : : : : : : : : : : : :
576 AGCAAGCAATTTATCTCATGACTTGACGGTGAATTAATGTATCTCGAA 625

61 heProGluValArgGluPhe.....AsnSerLeuProAsnAsnLysAla 75
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
626 ATGTAACAATTAATCAACACGACGACGAAACATTTGAATCTGCAAGGCT 675

76 SerSerAspThr.....AlaSerLeuAsnLysGlnHisAspAl 88
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
676 AGCAGCGATTCTTATTTGGAATGTCTACTCTTTTAATTTGAGAGAAGATTC 725

88 aAspPheLys.....LysTyrIleLysLeuLeuIleAsnAsnAsp 101
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
726 AAGTTTACCTTTATCAATACGTTAACTCTGCCAGAAATGGTGAT 771

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAAT58840
seq_documentation_block:
ID AAT58840 standard; DNA; 580073 BP.
XX
AC
AT58840;
XX
DT 27-MAR-1997 (first entry)
XX
DE Mycoplasma genitalium genome.
XX
KW M. genitalium; DNAA; DNA gyrase; origin of replication;
KW megabase shotgun sequencing method; open reading frame; ORF; ss.
XX
OS Mycoplasma genitalium.
XX
FH Key Location/Qualifiers
CDS 8552..9184
FT /*tag= a
FT /label= MG006
FT /note= "Previously identified as MORF-20076, the
FT encoded protein shows 27.59 percentage
FT identity to thymidylate kinase (CDC8)
FT from Saccharomyces cerevisiae"
FT
FT CDS 11252..12040
FT /*tag= b
FT /label= MG009
FT /note= "Previously identified as MORF-20078, the
FT encoded protein shows 35.43 percentage
FT identity to the Bacillus subtilis hypothetical
FT protein covered in accession number
FT GB:D26185_102"
FT
FT CDS 12069..12725
FT /*tag= c
FT /label= MG010
FT /note= "Previously identified as MORF-20079, the
FT encoded protein shows 25.73 percentage
FT identity to DNA primase (dnaE) from
FT Clostridium acetobutylicum"
FT
FT CDS complement (13570..14247)
FT /*tag= d
FT /label= MG012
FT

/note= "Previously identified as MORF-20080, the
FT encoded protein shows 31.50 percentage
FT identity to the ribosomal protein S6
FT modification protein (rimK) from Escherichia
FT coli"
FT complement (14396..15217)
FT /*tag= e
FT /label= MG013
FT /note= "Previously identified as MORF-19823, MORF-20080
FT and MORF-20081, the encoded protein shows 33.04
FT percentage identity to 5,10-methylene-tetra-
FT hydrofolate dehydrogenase (fold) from E. coli"
FT
FT 17474..19243
FT /*tag= f
FT /label= MG015
FT /note= "Previously identified as MORF-20084, the
FT encoded protein shows 32.23 percentage
FT identity to transport ATP-binding protein
FT (msbA) from E. coli"
FT
FT 26478..27344
FT /*tag= g
FT /label= MG023
FT /note= "Previously identified as MORF-20092, the
FT encoded protein shows 45.96 percentage
FT identity to fructose-bisphosphate aldolase
FT (tsr) from B. subtilis"
FT
FT 27345..28448
FT /*tag= h
FT /label= MG024
FT /note= "Previously identified as MORF-19826 and
FT MORF-20093, the encoded protein shows 46.84
FT percentage identity to GTP-binding protein
FT from E. coli"
FT
FT 36987..38978
FT /*tag= i
FT /label= MG032
FT /note= "Previously identified as MORF-20099, the
FT encoded protein shows 26.82 percentage
FT identity to ATP-dependent nuclease (addA)
FT from B. subtilis"
FT
FT 39242..39904
FT /*tag= j
FT /label= MG033
FT /note= "Previously identified as MORF-20100, the
FT encoded protein shows 35.90 percentage
FT identity to glycerol uptake facilitator
FT (glpF) from B. subtilis"
FT
FT complement (39873..40514)
FT /*tag= k
FT /label= MG034
FT /note= "Previously identified as MORF-20101, the
FT encoded protein shows 48.13 percentage
FT identity to thymidylate kinase (tdk)
FT from B. subtilis"
FT
FT 40543..41787
FT /*tag= l
FT /label= MG035
FT /note= "Previously identified as MORF-20102, the
FT encoded protein shows 30.71 percentage
FT identity to histidyl-trNA synthetase (hiss)
FT from Mycobacterium leprae"
FT
FT complement (44751..46277)
FT /*tag= m
FT /label= MG038
FT /note= "Previously identified as MORF-20105, the
FT encoded protein shows 46.83 percentage
FT identity to glycerol kinase (glpK)
FT from E. coli"
FT
FT complement (46268..47422)
FT /*tag= n
FT /label= MG039
FT /note= "Previously identified as MORF-19831 and
FT MORF-20106, the encoded protein shows 43.20
```



```
FT FT percentage identity to glycerol-3-phosphate
FT FT dehydrogenase (GUT2) from S. cerevisiae"
FT CDS 49377..49643
FT FT /*tag= O
FT FT /label= MG041
FT FT /note= "The encoded protein shows 48.86 percentage
FT FT identity to phosphohistidinoprotein-hexose
FT FT phosphotransferase (ptsH) from Mycoplasma
FT FT capricolum"
FT CDS 50060..51520
FT FT /*tag= P
FT FT /label= MG042
FT FT /note= "Previously identified as MORF-19832 and
FT FT MORF-20108, the encoded protein shows 41.92
FT FT percentage identity to spermidine/
FT FT putrescine transport ATP-binding protein
FT FT (potA) from E. coli"
FT CDS 51525..52382
FT FT /*tag= q
FT FT /label= MG043
FT FT /note= "Previously identified as MORF-20110, the
FT FT encoded protein shows 26.51 percentage
FT FT identity to spermidine/putrescine transport
FT FT system permease protein (potB) from E. coli"
FT CDS 52366..53220
FT FT /*tag= r
FT FT /label= MG044
FT FT /note= "Previously identified as MORF-20111, the
FT FT encoded protein shows 29.45 percentage
FT FT identity to spermidine/putrescine transport
FT FT system permease protein C (potC) from E. coli"
FT CDS 54658..55605
FT FT /*tag= s
FT FT /label= MG046
FT FT /note= "Previously identified as MORF-20112, the
FT FT encoded protein shows 36.60 percentage
FT FT identity to sialoglycoprotease (gcp)
FT FT from Pasteurella haemolytica"
FT CDS complement (56970..58310)
FT FT /*tag= t
FT FT /label= MG048
FT FT /note= "Previously identified as MORF-19834,
FT FT MORF-20114 and MORF-20115, the encoded protein
FT FT shows 43.02 percentage identity to signal
FT FT recognition particle protein (fth) from B.
FT FT subtilis"
FT CDS 58117..59079
FT FT /*tag= u
FT FT /label= MG049
FT FT /note= "Previously identified as MORF-20114 and
FT FT MORF-20115, the encoded protein shows 44.78
FT FT percentage identity to purine-nucleoside
FT FT phosphorylase (deob) from E. coli"
FT CDS 59083..59754
FT FT /*tag= v
FT FT /label= MG050
FT FT /note= "Previously identified as MORF-20117, the
FT FT encoded protein shows 83.03 percentage
FT FT identity to deoxyribose-phosphate aldolase
FT FT (deoc) from Mycoplasma pneumoniae"
FT CDS complement (64898..65731)
FT FT /*tag= w
FT FT /label= MG056
FT FT /note= "Previously identified as MORF-20122, the
FT FT encoded protein shows 30.25 percent
FT FT identity to the protein disclosed in
FT FT GB:D26185_99 from B. subtilis"
FT CDS complement (65713..66249)
FT FT /*tag= x
FT FT /label= MG057
FT FT /note= "Previously identified as MORF-20123, the
FT FT encoded protein shows 38.90 percentage
FT FT identity to the protein disclosed in
```

```
FT FT GB:D26185_104 from B. subtilis"
FT CDS 81047..82597
FT FT /*tag= y
FT FT /label= MG067
FT FT /note= "Previously identified as MORF-19845, the
FT FT encoded protein shows 28.84 percentage
FT FT identity to glutamic acid specific protease
FT FT (SPase) from Staphylococcus aureus"
FT CDS 91065..91919
FT FT /*tag= z
FT FT /label= MG070
FT FT /note= "Previously identified as MORF-20136, the
FT FT encoded protein shows 34.8 percentage
FT FT identity to ribosomal protein S2 (rps2)
FT FT from Spirulina plantensis"
FT alignment_scores:
FT Quality: 72.00 Length: 129
FT Ratio: 1.180 Gaps: 6
FT Percent Similarity: 47.287 Percent Identity: 27.132
FT alignment_block:
FT US-09-528-682-4 x AAT58840/rev ..
FT Align seg 1/1 to reverse of: AAT58840 from: 1 to: 580073
FT
FT 7 IleGlnAspAlaTyrGluProIleAlaAsnThrAsnThrThrGlnPh 23
FT ||||| ||||| ||||| ||||| ||||| |||||
FT 535817 ATTTTGTATGCACATCCAAATCTTAATGATCTCTACTATCAGCA 535768
FT
FT 23 eLeu.....AsnMetGlyAsnGluValAlaLeuAspG 34
FT ||||| ||||| ||||| ||||| |||||
FT 535767 ATTAGAACAACAACATAATAAAACATCCCTGTTATTGCTGATA 535718
FT
FT 34 lyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAspArg 50
FT :: ::::: ||||| ::::: |||||
FT 535717 ATGTTTTTCAACAACAGCTTTAATGTTTAGTGAA.....ATGGATGAT 535674
FT
FT 51 AspLeuPheGlySerAlaAsnIleAsp..... 59
FT ||| ::|
FT 535673 AGTATTTTAAAGAAGCGCCAGTGATATATTAGATTACACCAAGGCT 535624
FT
FT 60 ....GlyPheProGluValArgGluPheAsnSerLeu..... 70
FT ::::: ||||| ::::: |||||
FT 535623 CTTAAGTTATTAACTGGAGTAAACTTTAATGATTTAATCAGGATTAAAA 535574
FT
FT 71 .....ProAsnAsnLysAlaSerSerAspThrAlaSer 81
FT ||||| ::| ||||| ||||| |||||
FT 535573 GTGATGTCATTATAGTAGCTAATGATTTAACACCAAGTCAACAGCAACT 535524
FT
FT 82 LeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLeuLeu1 98
FT ||||| ||||| ||||| ||||| |||||
FT 535523 CTTAAT.....AAAAGTATGTCAG..... 535503
FT
FT 98 eAsnAsnAspGlyPhePheSerAsnAsnGlyGlyLys 110
FT ||||| ::::: |||||
FT 535502 .....GCTTTTAACTGAGAGTGGTGTAA 535476
FT
FT seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAC90510
FT seq_documentation_block:
FT ID AAC90510 standard; cdna; 7492 BP.
FT XX
FT AC AAC90510;
FT XX
FT DT 15-MAR-2001 (first entry)
FT XX
FT DE Mouse factor VIII CDNA.
FT XX
FT KW Mouse; factor VIII; fviii; coagulant; clotting factor; haemophilic;
FT KW Immune detection evasion; ss.
FT XX
FT OS Mus musculus.
```











3793 GGTCTTAAAGACCAATGGAATAAATTCCTTGAACCTCTGAGCAAGAACAT 3842  
 89 Asp.....PheLysLysTyrIle..... 94  
 3843 AGTCCAAAGCAATTAGTATATTTAATGCTTTAAAAAATATCTAAAAATCA 3892  
 95 .....LysLeuLeuIleAsnAspGlyPheP 104  
 3893 AAGTTTCTGTGACAGAGAAAAATAAGTCACAGTAGAACAGGATGGATTTA 3942  
 104 heSerAsnAsnGlyGlyLys 110  
 3943 CAAAGAACATAGGACTTAAA 3962

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAx91164

seq\_documentation\_block:  
 ID AAX91164 standard; cDNA to mRNA; 7493 BP.

XX AC AAX91164;  
 XX DT 26-NOV-1999 (first entry)  
 XX DE Mouse factor VIII protein encoding nucleotide sequence.  
 XX Factor VIII protein; hybrid protein; porcine; mouse; immunogenicity;  
 KW antigenic; procoagulant; factor X activation; haemophilia; human; ds.  
 XX OS Mus musculus.  
 XX PN W09946274-AL.  
 XX PD 16-SEP-1999.  
 XX PF 10-MAR-1999; 99WO-US05193.  
 XX PR 10-MAR-1998; 98US-0037601.  
 XX (UYEM-) UNIV EMORY.  
 XX PA Lollar JS;  
 XX PI WPI; 1999-551355/46.  
 XX DR P-PSDB; AAY31596.  
 XX PT New porcine and modified human factor VIII proteins for treating  
 XX hemophilia -  
 XX PS Examples; Page 131-136; 187pp; English.

XX The invention provides DNA encoding porcine factor VIII, hybrid porcine  
 CC /human factor VIII, or modified human factor VIII having reduced  
 CC immunogenicity. Active factor VIII increases catalytic efficiency of  
 CC factor IXa towards factor X activation. The factor VIII proteins of the  
 CC invention can be produced by standard recombinant methodology and have  
 CC less antigenic activity, or greater procoagulant activity, than prior art  
 CC factor VIII. The invention is used to treat haemophilia. The present  
 CC sequence represents the nucleotide sequence encoding the A and C domains  
 CC of mouse factor VIII protein.

XX SQ Sequence 7493 BP; 2487 A; 1503 C; 1436 G; 2067 T; 0 other;

alignment\_scores:  
 Quality: 71.50 Length: 140  
 Ratio: 0.993 Gaps: 5  
 Percent Similarity: 51.429 Percent Identity: 21.429

alignment\_block:  
 US-09-528-682-4 x AAX91164 ..

Align seg 1/1 to: AAX91164 from: 1 to: 7493

4 ThrValArgIleGlnAspAlaTyrGluProIle...AlaAsnThrAsnTh 19  
 3543 ACAACAGACTGTCAGATGCCATATATAAGTCATAGTGCAGATTCAAGA 3592  
 19 rThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgT 36  
 3593 AGTACACAGCTTTGATTCATGATGGAACACTTTTAGGCAAAAATTCACAT 3642  
 36 Yr.....SerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAspArg 50  
 3643 ATTGAGACTAAACCATATGCTAAATAGAACTACCTCAACAAAAATAAA 3692  
 51 AspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPh 67  
 3693 GACATATTTCATAGAAAAGATGAAGATCCTATATCCACAAGATGAAGAGAA 3742  
 67 eAsnSerLeuProAsnAsn..... 73  
 3743 TACAATCATGCCATTTTCCAGATGTTGTTCTGTGAGAACTTCAAAT 3792  
 74 .....LysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAla 88  
 3793 GGTTTAAAAAGACCAATGGAATAATTCCTTGAACCTCTGAGCAAGAACAT 3842  
 89 Asp.....PheLysLysTyrIle..... 94  
 3843 AGTCCAAAGCAATTAGTATATTTAATGCTTTAAAAAATATCTAAAAATCA 3892  
 95 .....LysLeuLeuIleAsnAsnAspGlyPheP 104  
 3893 AAGTTTCTGTGACAGAGAAAAATAAGTCACAGTAGAACAGGATGGATTTA 3942  
 104 heSerAsnAsnGlyGlyLys 110  
 3943 CAAAGAACATAGGACTTAAA 3962

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AA117717

seq\_documentation\_block:

ID AA117717 standard; DNA; 3300 BP.

XX AC AA117717;  
 XX DT 21-MAY-1996 (first entry)  
 XX DE Caulobacter rsaa gene.  
 XX KW S-layer; rsaa gene; surface protein; biofilm; vaccine; ds.  
 XX OS Caulobacter crescentus strain CB15.  
 XX Key Location/Qualifiers  
 FH -35\_signal 5..7  
 FT /tag= a  
 FT -10\_signal 28..31  
 FT /tag= b  
 FT RBS 88..91  
 FT /tag= c  
 FT CDS 101..3181  
 FT /tag= d  
 FT /transl\_except= pos:1229..1231;\_aa:Gly  
 FT /transl\_except= pos:2627..2629;\_aa:Gly  
 FT /transl\_except= pos:2630..2632;\_aa:Ser  
 FT 3214..3230  
 FT /tag= e  
 FT /note= "putative transcription terminator  
 FT palindromic"  
 FT terminator 3233..3248  
 FT /tag= f  
 FT /note= "putative transcription terminator  
 FT palindromic"  
 XX PN US5500353-A.



```

XX 19-MAR-1996.
XX 09-JUN-1992; 92US-0895367.
XX 09-FEB-1994; 94US-0194290.
XX 09-JUN-1992; 92US-0895367.
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Bingle WH, Smit J;
XX WPI; 1996-171046/17.
XX P-PSDB; AAR94014.
XX
XX Expression and presentation of polypeptide heterologous to
XX Caulobacter S-layer protein to bacterium's environment - by cloning
XX in-frame into Caulobacter rsaa gene, and expressing as fusion prod.
XX with S-layer protein
XX
XX Example 2; Fig 6a-c; 22pp; English.
XX
XX The rsaa gene (AAT17717) of Caulobacter crescentus CB15 codes for
XX the paracrystalline S-layer protein (AAR94014). Restriction
XX sites within the gene permit the insertion of heterologous DNA.
XX The encoded polypeptide is then expressed and presented on the
XX surface of the Caulobacter as a fusion protein with the S-layer
XX protein. The bacterium is cultured as a biofilm in a bioreactor
XX or may be used to present an antigenic epitope (see AAR94016 and
XX (AAR94016) to the environment e.g. for use as a vaccine.
XX
XX Sequence 3300 BP; 548 A; 1218 C; 1005 G; 529 T; 0 other;
SQ

```

alignment\_scores:

Quality:	71.00	Length:	112
Ratio:	1.164	Gaps:	5
Percent Similarity:	54.464	Percent Identity:	28.571

alignment\_block:

US-09-528-682-4 x AAT17717 ..

Align seg 1/1 to: AAT17717 from: 1 to: 3300

```

1 AspTyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 GACTACCTGACCGCCTCGTGGCGC.....GCCAACAC 639

17 rAsn...ThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
640 GCCGTTACGGCGCGTCCGACATCGATCGCGCTCAAGCGCGCCTGA 689

33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGly..... 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
690 TCGGCACCATCTCTGAACGCGCGCATCGGTGCGGCATCGGTGCTACGCG 739

48 .....MetAspArgAspLeuPhe....GlySerAlaAsnIle 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
740 ACCGCGACGCGCGCGATGATCAACGACCTCTCGGACGCGCGCTGCGAC 789

58 eAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysA 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
790 CGACACGCGCGTGGGTGAACCTGTTCACCGCCTATCCCTGCTGGGCG 839

75 laSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLys 91
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
840 TGTCGGGTTCGACCTCTCGCTGACACCGCGCACCGACAC..... 880

92 LysTyrIleIleLysLeuLeuIleAsnAsnAspGlyPhe 103
::: ||||||| |||
881 .....CTGACGGGACCGCGCAACACGACGCTTC 910

```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAV01866

```

seq_documentation_block:
ID AAV01866 standard; DNA; 3300 BP.
XX
XX AC AAV01866;
XX
XX DT 20-APR-1998 (first entry)
XX
XX DE Caulobacter crescentus S-layer rsaa gene.
XX
XX KW S-layer; rsaa gene; Caulobacter; vaccine; antigenic; ligand; enzyme;
XX metallothionein; heavy metal; water; sewage; xylanase; cellulase;
XX wood pulping; ss.
XX
XX OS Caulobacter crescentus.
XX
XX FH Key Location/Qualifiers
XX FT -35_signal 5..7
XX FT /*tag= a
XX FT -10_signal 28..31
XX FT /*tag= b
XX FT RBS 3292..91
XX FT /*tag= c
XX FT /standard_name= Shine_Dalgarno_sequence
XX FT CDS 101..3181
XX FT /*tag= d
XX FT /product= "S-layer protein"
XX
XX PN WO9734000-A1.
XX
XX PD 18-SEP-1997.
XX
XX PF 10-MAR-1997; 97WO-CA00167.
XX
XX PR 12-MAR-1996; 96US-0614377.
XX
XX PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX PI Bingle WH, Nomellini JF, Smit J;
XX
XX DR WPI; 1997-470880/43.
XX
XX DR P-PSDB; AAW37490.
XX
XX
XX PT New DNA containing sequence for C-terminal region of Caulobacter
XX S-layer protein - expressed as fusion proteins containing antigenic
XX peptides in Caulobacter, useful as live vaccines
XX
XX PS Example 2; Fig 6; 58pp; English.
XX
XX CC The present sequence encodes a Caulobacter S-layer protein used in an
XX example of the present invention. A new DNA construct has been developed
XX which contains at least one restriction site for insertion of DNA
XX upstream of DNA encoding a C-terminal region of at least the last 82
XX amino acids (aa) of Caulobacter S-layer protein. Caulobacter containing
XX the DNA constructs above additionally containing a sequence encoding a
XX heterologous polypeptide, are particularly useful in live vaccines
XX (where the heterologous polypeptide is an antigen). They can also be
XX used for production of e.g. ligands, enzymes or other proteins, e.g.
XX metallothioneins to remove heavy metals from water or sewage, or
XX xylanase or cellulase for use in wood pulping. All known Caulobacter
XX strains are harmless, and stable in outdoor environments, including
XX water (so suitable for vaccinating fish) or soil. They are well suited
XX for growing in biofilm reactors and produce S-layer proteins, which is
XX an ideal system for presentation of antigens, at high level.
XX
XX SQ Sequence 3300 BP; 548 A; 1218 C; 1005 G; 529 T; 0 other;

```

alignment\_scores:

Quality:	71.00	Length:	112
Ratio:	1.164	Gaps:	5
Percent Similarity:	54.464	Percent Identity:	28.571



```
alignment_block:
US-09-528-682-4 x AAV01866  ..
Align seg 1/1 to: AAV01866 from: 1 to: 3300
1 AsPTyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
|||||:|||||:  :|||:  |||||
608 GACTACCTGACCGCTTCGTGCGC.....GCCAACAC 639
17 rAsn...ThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuA 33
| :|||:|||||:  :|||:  |||||
640 GCCGTTACGGCCGCTGCGGACATCGATCTGCGCGTCAAGGCCGCTGA 689
33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGly..... 47
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
690 TCGGCACCATCTGAACGCCGCCGCTGCGGACATCGGTGTTACGCG 739
48 .....MetAspArgAspLeuPhe...GlySerAlaAsnII 58
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
740 ACCGCCACGGCGGATGATCAGCAGCTGTGCGGACGGCGCTGTGCGAC 789
58 eAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysA 75
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
790 CGACAACGGCGCTGGGTGAACCTGTTCACGCCCTATCCGTGCGGGCG 839
75 laSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLys 91
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
840 TGTGCGGTTCGACCTCTCGCTGACCGCGGCCGACGACCC..... 880
92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103
::: ||||| |||
881 .....CTGACGGGCACCGCCCAACAGCACGTC 910
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ50079

```
seq_documentation_block:
ID AAZ50079 standard; DNA; 3300 BP.
XX
AC AAZ50079;
XX
DT 04-MAY-2000 (first entry)
XX
DE Caulobacter crescentus surface layer protein encoding gene.
XX
KW Surface layer protein; S-layer secretion signal; antibiotic; vaccine;
KW recombinant fusion protein cleavage; enzyme; protein polymer;
KW antibacterial enzyme; foodstuff; ds.
XX
OS Caulobacter crescentus.
XX
FH Key Location/Qualifiers
FT CDS 101..3181
FT FT /*tag= a
FT FT /product= "S-layer protein"
XX
PN W0200004170-A1.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-CA00637.
XX
PR 14-JUL-1998; 98CA-2237704.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI Smit J;
XX
WPi; 2000-182434/16.
DR P-PSDB; AAY44757.
XX
PT Cleavage of Caulobacter produced recombinant fusion proteins useful for
XX producing vaccine peptides
XX
```

```
PS Example 1; Pages 19-20; 33pp; English.
XX
CC The patent discloses a method for cleaving a recombinant fusion protein
CC which is produced by Caulobacter and consists of Caulobacter surface
CC layer (S-layer) protein (containing the C-terminal secretion signal) and
CC a target protein heterologous to Caulobacter. The cleavage of target
CC protein from the S-layer protein is carried out under mild acid
CC conditions so that cleavage occurs at aspartate-proline dipeptide site
CC without solubilising the protein. The cleavage is accomplished while the
CC fusion protein is in an insoluble aggregate form which facilitates
CC purification of the protein. The method is useful for producing pure
CC proteins including recombinant human and animal therapeutic antibiotic
CC and vaccine peptides, enzymes, protein polymers, and antibacterial
CC enzymes for foodstuffs.
CC The present sequence is a gene encoding C. crescentus S-layer protein.
CC The S-layer secretion signal, corresponding to the C-terminal portion of
CC the protein from amino acid 690 onwards, is fused with a target sequence
CC for construction of a recombinant fusion construct which is expressed in
CC Caulobacter.
XX
SQ Sequence 3300 BP; 548 A; 1218 C; 1005 G; 529 T; 0 other;

alignment_scores:
Quality: 71.00 Length: 112
Ratio: 1.164 Gaps: 5
Percent Similarity: 54.464 Percent Identity: 28.571

alignment_block:
US-09-528-682-4 x AAZ50079  ..
Align seg 1/1 to: AAZ50079 from: 1 to: 3300
1 AsPTyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
|||||:|||||:  :|||:  |||||
608 GACTACCTGACCGCTTCGTGCGC.....GCCAACAC 639
17 rAsn...ThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuA 33
| :|||:|||||:  :|||:  |||||
640 GCCGTTACGGCGCTGCGGACATCGATCTGCGCGTCAAGGCCGCTGA 689
33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGly..... 47
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
690 TCGGCACCATCTGAACGCCGCCGCTGCGGACATCGGTGTTACGCG 739
48 .....MetAspArgAspLeuPhe...GlySerAlaAsnII 58
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
740 ACCGCCACGGCGGATGATCAGCAGCTGTGCGGACGGCGCTGTGCGAC 789
58 eAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysA 75
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
790 CGACAACGGCGCTGGGTGAACCTGTTCACGCCCTATCCGTGCGGGCG 839
75 laSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLys 91
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
840 TGTGCGGTTCGACCTCTCGCTGACCGCGGCCGACGACCC..... 880
92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103
::: ||||| |||
881 .....CTGACGGGCACCGCCCAACAGCACGTC 910
92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103
::: ||||| |||
881 .....CTGACGGGCACCGCCCAACAGCACGTC 910
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA552160

```
seq_documentation_block:
ID AAS52160 standard; DNA; 3054 BP.
XX
AC AAS52160;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #577.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
```







CC programmes. The antisense nucleic acid sequence is also useful to screen.  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence encodes an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

XX Sequence 3057 BP; 1244 A; 491 C; 619 G; 703 T; 0 other;

alignment\_scores:  
 Quality: 70.50 Length: 105  
 Ratio: 1.282 Gaps: 3  
 Percent Similarity: 52.381 Percent Identity: 25.714

alignment\_block:

US-09-528-682-4 x AAS55103 ..

Align seg 1/1 to: AAS55103 from: 1 to: 3057

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31

1036 GCCAATTTAAATGATCGATTGAGACATTTAATAAAGCGAAT..... 1077

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyM 48

1078 .....AATAGATTTTCGCATGTTGCATTTTAAACCTAATAATGGTA 1120

48 etAspArgAsp.....LeuPheGlySerAlaAsn 57

1121 AAACGACAAAGTGTGACTGTTACTGGAACCTTTAATGAAAGGTAGTAATCAG 1170

58 IleAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLy 74

1171 AATGGAATAACCAAGAAAGTAGGATTTTGAATACTTGGGTAATAATGA 1220

74 sAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheL 91

1221 AGACATACGCAAGAGTGTATATGCAATACGACAGACATACCTCTTAATTTA 1270

91 ysLysTyrIle.....LysLeuLeuIleAsnAsnAspGly 102

1271 AAGAAGTCACAAAGTAATATGATGGGAATTTGAATTTACAAAATAATGGA 1320

103 PhePheSerAsnAsn 107

1321 AGCTATTTCATTGAAT 1335

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: AAN82053

seq\_documentation\_block:

ID AAN82053 standard; DNA; 3280 BP.

XX AC AAN82053;

XX 05-JAN-1990 (first entry)

XX DNA encoding fibronectin binding protein.

XX Fibronectin binding protein; S.aureus; immunisation;

KW mastitis; ruminants; wound infection; diagnosis.

XX Staphylococcus aureus.

XX Key Location/Qualifiers

FT CDS 57..313

FT /\*tag= a

PN EP294349-A.

XX 07-DEC-1988.

XX

XX 30-MAY-1988; 88EP-0850188.

XX 01-JUN-1987; 87SE-0002272.

XX (ALFA ) ALFA-LAVAL AGRI INTERNATIONAL AB.

XX Hook M, Lindberg MK, Signas LC, Wadstrom TM, Froman G;

XX WPI: 1988-347978/49.

XX P-PSDB; AAP82115.

XX Hybrid DNA encoding Staphylococcus aureus fibronectin binding protein -

XX useful for immunisation and topical application to prevent

XX staphylococcal infections.

XX Disclosure; Fig. 8A:1-8A:2; 23pp; English.

XX The DNA encodes a S.aureus fibronectin binding protein which may be

XX included in a hybrid protein. The protein may be used to immunise

XX ruminants against staphylococcal mastitis, pref. when given at

XX 0.5-5 microg/kg, using 3 doses at 1-3 week intervals; and for

XX topical applicn. to prevent wound infection, pref. using an

XX isoconic saline soln. of concn. 25-250 microg/mg. The DNA and the

XX hybrid protein may be immobilised on a carrier and used to diagnose

XX staphylococcal infections.

XX Sequence 3280 BP; 1327 A; 509 C; 652 G; 792 T; 0 other;

XX

alignment\_scores:

Quality: 70.50 Length: 105

Ratio: 1.282 Gaps: 3

Percent Similarity: 52.381 Percent Identity: 25.714

alignment\_block:

US-09-528-682-4 x AAN82053 ..

Align seg 1/1 to: AAN82053 from: 1 to: 3280

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31

1092 GCCAATTTAAATGATCGATTGAGACATTTAATAAAGCGAAT..... 1133

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyM 48

1134 .....AATAGATTTTCGCATGTTGCATTTTAAACCTAATAATGGTA 1176

48 etAspArgAsp.....LeuPheGlySerAlaAsn 57

1177 AAACGACAAAGTGTGACTGTTACTGGAACCTTTAATGAAAGGTAGTAATCAG 1226

58 IleAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLy 74

1227 AATGGAATAACCAAGAAAGTAGGATTTTGAATACTTGGGTAATAATGA 1276

74 sAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheL 91

1277 AGACATACGCAAGAGTGTATATGCAATACGACAGACATACCTCTTAATTTA 1326

91 ysLysTyrIle.....LysLeuLeuIleAsnAsnAspGly 102

1327 AAGAAGTCACAAAGTAATATGATGGGAATTTGAATTTACAAAATAATGGA 1376

103 PhePheSerAsnAsn 107

1377 AGCTATTTCATTGAAT 1391

seq\_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: AAV74365

seq\_documentation\_block:

ID AAV74365 standard; DNA; 13740 BP.

XX























DR WPI: 2000-293136/25.  
 DR P-PSDB; AAY92177.

XX New cardiovascular system associated protein tyrosine phosphatases are  
 PT used in the diagnosis and treatment of e.g. immune disorders,  
 PT anti-proliferative disorders, metabolic disorders and cardiovascular  
 PT disorders

XX Claim 1; Page 155-156; 156pp; English.

XX This sequence encodes a cardiovascular system associated protein tyrosine  
 CC phosphatase 4 (CSA4PTP). The CSA4PTP nucleic acid and protein molecules  
 CC are used to modulate regulation of cellular processes. CSA4PTP nucleic  
 CC acid and protein molecules and modulators of CSA4PTP activity and  
 CC expression can be used to treat a subject with a disorder characterized  
 CC by aberrant CSA4PTP expression or activity. These disorders can include  
 CC an immune disorder, an anti-proliferative disorder, a proliferative  
 CC disorder e.g. renal and lung carcinomas, a metabolic disorder e.g.  
 CC diabetes, viral pathogenesis, a neural disorder, a cardiovascular  
 CC disorder e.g. hypertension or coronary heart disease or a disorder rising  
 CC from improper phosphorylation of a phosphorylated protein. Anti-CSA4PTP  
 CC antibodies are used to isolate CSA4PTP by standard techniques, to  
 CC facilitate the purification of natural and recombinantly produced CSA4PTP  
 CC from cells, to detect CSA4PTP protein in cell supernatant for evaluating  
 CC the level of CSA4PTP expression and to monitor protein levels in tissue as  
 CC part of a clinical testing procedure to determine efficacy of a treatment  
 CC procedure. CSA4PTP nucleic acids, antibodies and protein molecules can be  
 CC used in screening assays and in predictive medicine e.g. prognostic  
 CC assays, monitoring clinical trials and pharmacogenetics.

XX SQ Sequence 663 BP; 126 A; 222 C; 182 G; 133 T; 0 other;

alignment\_scores:  
 Quality: 69.50 Length: 70  
 Ratio: 1.448 Gaps: 3  
 Percent Similarity: 68.571 Percent Identity: 30.000

alignment\_block:

US-09-528-682-4 x AAA09032/rev ..

Align seg 1/1 to reverse of: AAA09032 from: 1 to: 663

32 LeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMe 48  
 ||||| ||| :  
 567 CTGGATGGCCTCTACCGCGTCTCTCATGATCATGAGGAGGCCA 518  
 48 tAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGlu.... 63  
 :||| :  
 517 GGACAAG.....TGTGGCAGAGCG.....GCTTACCCCATGGCA 483  
 64 ..ValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspThr 79  
 ||| :  
 482 CAGTGTACACGACCGCGCTTGGGGAACACTGAGGCGAGCTCGGATGTA 433  
 80 AlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLe 96  
 :||| :  
 432 TCGACCAACAGCAGCAAGTAGACACTGAGTCTGAGGAGGGTGTGCT 383  
 96 uLeuIleAsn 99  
 :  
 382 CGGCCTCAAT 373

seq\_name: /SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF63571

seq\_documentation\_block:

ID AAF63571 standard; cDNA; 904 BP.

XX AAF63571;

XX 11-MAY-2001 (first entry)

DE Human phosphatase LOC51207\_h coding sequence.

XX

KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
 KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
 KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
 KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
 KW schizophrenia; hamartoma; ss.

XX Homo sapiens.

XX WO200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGF-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI: 2001-211226/21.

XX P-PSDB; AAB73219.

XX New protein phosphatase polypeptide for diagnosing and treating  
 PT phosphatase related disorders such as cancer, schizophrenia, cardiac  
 PT dysfunction and/or vascular disorders

XX Claim 1; Fig 4; 138pp; English.

XX The present invention relates to phosphatase proteins and coding  
 CC sequences. The present sequence is one such phosphatase coding sequence.  
 CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins  
 CC modified by phosphorylation of serine, threonine or tyrosine residues.  
 CC The phosphatases are useful for treating a variety of diseases: for  
 CC example cancer e.g. breast, urogenital, prostate, head, neck, lung  
 CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung  
 CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach  
 CC cancer, glioblastoma, colorectal cancer and thyroid cancer.  
 CC pathophysiological hypoxia, cardiac dysfunction and/or vascular  
 CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre  
 CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,  
 CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and  
 CC hamartomas.

XX SQ Sequence 904 BP; 183 A; 282 C; 249 G; 190 T; 0 other;

alignment\_scores:

Quality: 69.50 Length: 70  
 Ratio: 1.448 Gaps: 3  
 Percent Similarity: 68.571 Percent Identity: 30.000

alignment\_block:

US-09-528-682-4 x AAF63571/rev ..

Align seg 1/1 to reverse of: AAF63571 from: 1 to: 904

32 LeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMe 48  
 ||||| ||| :  
 568 CTGGATGGCCTCTACCGCGTCTCTCATGATCATGAGGAGGCCA 519  
 48 tAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGlu.... 63  
 :||| :  
 518 GGACAAG.....TGTGGCAGAGCG.....GCTTACCCCATGGCA 484  
 64 ..ValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspThr 79  
 ||| :  
 483 CAGTGTACACGACCGCGCTTGGGGAACACTGAGGCGAGCTCGGATGTA 434  
 80 AlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLe 96  
 :||| :  
 433 TCGACCAACAGGCAAGTAGACACTGAGTCTGAGGAGGGTGTGCT 384















CC syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian  
 CC tumours, uterine fibroids, autoimmune disorders, ectopic pregnancy,  
 CC teratogenesis, breast cancer, fibrocystic breast disease,  
 CC galactorrhea, disruptions of spermatogenesis, abnormal sperm  
 CC physiology, testis cancer, prostate cancer, benign prostatic  
 CC hyperplasia, prostatitis, Peyronie's disease, male breast carcinoma  
 CC and gynecomastia; gastrointestinal disorders including, but are not  
 CC limited to, dysphagia, peptic oesophagitis, oesophageal spasm and  
 CC stricture, oesophageal carcinoma, dyspepsia, indigestion, gastritis,  
 CC gastric carcinoma, anorexia, nausea, emesis, gastroparesis, intestinal  
 CC tract infection, peptic ulcer, colitis, Whipple's disease, Mallory-Weiss  
 CC syndrome, irritable bowel syndrome, short bowel syndrome, diarrhoea,  
 CC constipation, cirrhosis, jaundice, and hepatic vein thrombosis; nervous  
 CC disorders including, but are not limited to, Alzheimer's disease,  
 CC amnesia, bipolar disorder, catatonia, cerebral neoplasms, Down's  
 CC syndrome, and dystonias; smooth muscle cell disorders including, but  
 CC not limited to, angina, anaphylactic shock, arrhythmia, cardiovascular  
 CC shock, migraine, and pheochromocytoma; musculoskeletal disorders  
 CC including muscular dystrophy, central core disease, nemaline myopathy,  
 CC centronuclear myopathy, lipid myopathy, inclusion body myositis,  
 CC thyrotoxic myopathy, and ethanol myopathy; immunological disorders  
 CC including AIDS, Addison's disease, adult respiratory distress  
 CC syndrome, allergy, ankylosing spondylitis, amyloidosis, anaemia,  
 CC asthma, atherosclerosis, autoimmune hemolytic anaemia, autoimmune  
 CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's  
 CC disease, atopic dermatitis, dermatomyositis, diabetes mellitus,  
 CC emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis  
 CC fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis,  
 CC Goodpasture's syndrome, gout, Grave's disease, Hashimoto's thyroiditis,  
 CC hypereosinophilia, irritable bowel syndrome, multiple sclerosis,  
 CC myasthenia gravis, myocardial or pericardial inflammation,  
 CC osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis,  
 CC Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjogren's  
 CC syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic  
 CC sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis,  
 CC Werner syndrome, complications of cancer, haemodialysis, and  
 CC extracorporeal circulation; viral, fungal, bacterial or protozoan  
 CC infections; trauma; and neoplastic disorders including adenocarcinoma,  
 CC leukemia, melanoma, myeloma, sarcoma, and various cancers. The REC  
 CC polynucleotide is a source of probes and primers which bind may be used  
 CC to detect REC in a sample from a patient (claimed). They may also be  
 CC administered as part of a gene therapy regime.

XX  
 SQ Sequence 1943 BP; 515 A; 537 C; 473 G; 418 T; 0 other;

alignment\_scores:  
 Quality: 69.00 Length: 72  
 Ratio: 1.683 Gaps: 4  
 Percent Similarity: 56.944 Percent Identity: 31.944

alignment\_block:  
 US-09-528-682-4 x AA234611 ..

Align seg 1/1 to: AA234611 from: 1 to: 1943

13 ProfileAlaAsnThrThrThrGlnPheLeuAsnMetGlyAsnG1 29  
 ||| .....  
 863 CCCGAGATTCACAAATTCAGTGACACATCTACACATGATTATCC 912  
 ||| .....  
 29 uValalaLeuAspGlyArgTyrSerAsnTyrAlaLeu..... 41  
 ||| .....  
 913 AGTGGCGGCTGGAGGCGCGGTCCAACTCCCGATGGTCCCGGCTCGG 962  
 ||| .....  
 42 .....IleSerAlaGluGlyMetAsp.....ArgAsp 51  
 .....  
 963 ACGGTCCGATGGCGGATGGTGGCATGGAGCCACACCATGAATGGA 1012  
 .....  
 52 LeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheAs 68  
 .....  
 1013 TCATTAGGCTGAGGACATAGACGACTCCAAAA.....AA 1050  
 .....  
 68 nSerLeuProAsnAsn 73

|||||  
 1051 TTCT...CCTAACAC 1063

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-nbml/NA1996.DAT:AA08098

seq\_documentation\_block:  
 ID AAT08098 standard; DNA; 2278 BP.  
 AC AAT08098;  
 XX  
 DT 04-JUL-1996 (first entry)  
 XX  
 DE Enterohaemorrhagic E.coli hlyA gene.  
 KW Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;  
 KW detection; hlyA gene; probe; primer; enterohaemorrhagic colitis;  
 KW haemolytic uremic syndrome; mesenteric adenitis; ss.  
 XX  
 OS Escherichia coli (enterohaemorrhagic).  
 FH Key Location/Qualifiers  
 CDS 2..2278  
 FT /\*tag= a  
 FT /product= hlyA\_gene\_product  
 FT /note= "incomplete CDS; does not include start  
 codon"  
 XX  
 PN US5475098-A.  
 XX  
 PD 12-DEC-1995.  
 XX  
 PF 14-JUN-1994; 94US-0258188.  
 XX  
 PR 14-JUN-1994; 94US-0258188.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 XX Hall RH, Xu JG;  
 XX  
 XX WPI: 1996-048546/05.  
 DR P-PSDB; AAR86998.  
 XX  
 PT Enterohaemorrhagic E. coli (EHEC) nucleic acid sequences - useful  
 PT for probe and primer design for sensitive and specific detection of  
 PT EHEC  
 XX  
 PS Claim 6; Columns 31-38; 32pp; English.  
 XX  
 CC Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic  
 CC colitis, haemolytic uremic syndrome and mesenteric adenitis have  
 CC been found to carry a hlyA gene and a hlyB gene, separated by an  
 CC intergenic region. The hly genes and the intergenic region are  
 CC absent from bacteria not associated with these diseases and so  
 CC provide a useful target for detecting EHEC pathogens, esp. O157:H7  
 CC serotype E.coli. The present sequence is the EHEC hlyA gene.  
 XX  
 SQ Sequence 2278 BP; 745 A; 344 C; 550 G; 639 T; 0 other;

alignment\_scores:  
 Quality: 69.00 Length: 98  
 Ratio: 1.211 Gaps: 4  
 Percent Similarity: 58.163 Percent Identity: 25.510

alignment\_block:  
 US-09-528-682-4 x AAT08098 ..

Align seg 1/1 to: AAT08098 from: 1 to: 2278

23 PheLeuAsn.....MetGlyAsnGluValalaLeuAspGlyArgTyrSe 37  
 :|||||: .....  
 1745 TAICTGAATGCTGTTTGTGTAATGATATATTATGTTATGGCAAACTA 1794



















```

1008 GAGCATTGGCTGTATATACCGGATCCAGTTCATCCACCGATGACAC 1057
79 rlaSerLeuAnLys.....GlnHisAspAlaAspPheLysLysTyri 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1058 GCCTTCACGTGGACCGACATCTCTCATGGCAGTGTGTCTCTCTCC 1107
94 leLysLeuLeuIleAsnAn.....AspGlyPhe 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1108 AGATTTTAAAGCCAAACAGGTCAAGAGATCGCAAGCCCTTGATGATTC 1157
104 PheSerAsnAnGlyGly 109
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1158 TACAGCCATGGGATGGGA 1175

```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA569724

seq\_documentation\_block:

ID AA569724 standard; cDNA; 4290 BP.

XX AA569724;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #5528.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR P-PSDB; ABG05537.

XX New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -

PS Claim 1; SEQ ID No 5528; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AA564197-AA594564 represent novel human  
diagnostic coding sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 4290 BP; 1315 A; 944 C; 1137 G; 894 T; 0 other;

alignment\_scores:

Quality: 68.00 Length: 106  
Ratio: 1.172 Gaps: 6  
Percent Similarity: 54.717 Percent Identity: 29.245

alignment\_block:

US-09-528-682-4 x AA569724 ..

Align seg 1/1 to: AA569724 from: 1 to: 4290

```

23 PheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArg.....TyrSe 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 TTCTCTGGATGTTGGACTG.....GGCAGAGAGTGTACCTC 131
37 rAsnTyriAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 AAACAAGGTGTACTTAAAGAGATCTGGGAGTGTATCTGACCTCTTC 181
54 lySerAlaAsnIleAspGly.....PheProGluValArgGlu 66
|||||:|||||:|||||:|||||:|||||:|||||:
182 ACTCACCCAGTGTACATGGACGATCATCTTCCCAAGCCAGAGAA 231
67 .....PheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
|||||:|||||:|||||:|||||:|||||:|||||:
232 GAGCATTGGCCCTGTATATCATCCGGATCCAGTTCATCCACCGATGACAC 281
79 rlaSerLeuAnLys.....GlnHisAspAlaAspPheLysLysTyri 94
|||||:|||||:|||||:|||||:|||||:|||||:
282 GGCTTCACGTGGACCGACATCTCTCATGGCAGTGTGTCTCTCTCC 331
94 leLysLeuLeuIleAsnAn.....AspGlyPhe 103
|||||:|||||:|||||:|||||:|||||:|||||:
332 AGATTTTAAAGCCAAACAGGTCAAGAGATCGCAAGCCCTTGATGATTC 381
104 PheSerAsnAnGlyGly 109
:|||||:|||||:|||||:|||||:|||||:|||||:
382 TACAGCCATGGGATGGGA 399

```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA70204

seq\_documentation\_block:

ID AAA70204 standard; DNA; 4317 BP.

XX AAA70204;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:337.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;  
KW antimalarial; malaria; protozoacide; infection; insecticide; ds.  
XX Plasmodium falciparum.  
OS WO200025728-A2.  
XX 11-MAY-2000.  
XX 05-NOV-1999; 98WO-US26796.  
XX 05-NOV-1998; 98US-0107131.  
XX (HOFF/) HOFFMAN S.  
XX (CARU/) CARUCCI D.  
XX (GARD/) GARDNER M.  
XX (VENT/) VENTER J C.



PI Hoffman S, Carucci D, Gardner M, Venter JC;  
 XX WPI: 2000-365347/31.  
 XX  
 PT Proteins encoded by chromosome 2 of the human malarial parasite,  
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the  
 PT diagnosis of P.falciparum infection -  
 XX  
 XX  
 PS Disclosure; Page 530-531; 577pp; English.  
 XX  
 CC The present invention describes proteins and their fragments (I) encoded  
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.  
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)  
 CC vaccines against P. falciparum infection comprising (I) or (II).  
 CC (I) and (II) are useful for the development of vaccines against  
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal  
 CC antibody raised to immunogens comprising the sequences of (I), are  
 CC useful in the detection of infection with P. falciparum. Furthermore,  
 CC (I) (especially when they are rifins or secreted or membrane proteins)  
 CC can aid the identification of drugs to treat or prevent P. falciparum  
 CC infection, or they can be used to identify drug resistance in  
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the  
 CC subsequent identification of proteins encoded by it will help to expand  
 CC our understanding of parasite biology, a process hampered by the  
 CC complexity of the parasitic life cycle, and provide new targets for  
 CC vaccine and drug development. Parasite resistance to drugs and mosquito  
 CC resistance to insecticides have led to a resurgence of malaria in many  
 CC parts of the world, and there is a pressing need for vaccines and new  
 CC drugs. AAT70078 to AAT70287 and AAT8144 to AAT8352 represent nucleotide  
 CC and protein sequences given in the present invention, but which are not  
 CC specifically mentioned within the specification.  
 XX  
 SQ Sequence 4317 BP; 1862 A; 543 C; 582 G; 1330 T; 0 other;

alignment\_scores:  
 Quality: 68.00 Length: 55  
 Ratio: 1.943 Gaps: 2  
 Percent Similarity: 63.636 Percent Identity: 34.545

alignment\_block:

US-09-528-682-4 x AAT70204 ..

Align seg 1/1 to: AAT70204 from: 1 to: 4317

61 PheProGluValArgGluPheAsnSerLeuProAsnAsnLys.....AL 75  
 ||| .....:||||| ||| .....:|||||  
 3679 TTTCGAGATATCAAAATGTTGATCTCTCCACCAAGAAATAAATCAATC 3728  
 75 aserSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysL 92  
 : |||:|||||:||||| ||| |||:|||||  
 3729 CATATCTTCAACGGCAATTTGAATTATCAATGATAATCAATTTAACA 3778  
 92 ysTyr.....:|||||:|||||  
 || :||:|||||:|||||  
 3779 AAAGAGAAATGATATGGAACCAACAGTATATCTATGACTAATCAACGATCATCT 3828  
 103 PhePheSerAsnAsn 107  
 :||| |||  
 3829 ATGTTTTCGATGATAC 3843

seq\_name: /SIDSS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: AAX24044

seq\_documentation\_block:

ID AAX24044 standard: cDNA; 5276 BP.

XX

AC AAX24044;

XX

DT 30-JUN-1999 (first entry)

XX

DE Human brx cDNA.

XX

KW Brx protein; breast cancer; nuclear receptor-binding auxiliary protein;

KW anti-cancer; anti-proliferative; mitogen; transcription factor; human;  
 KW tumour suppressor; ovarian cancer; proliferative disorder; treatment;  
 KW immune tissue; prevention; reproductive tissue; ss.  
 XX Homo sapiens.  
 OS  
 XX WO9915544-A1.  
 PN  
 XX PD 01-APR-1999.  
 XX  
 XX PF 23-SEP-1998; 98WO-US19782.  
 XX  
 XX PR 23-SEP-1997; 97US-0059621.  
 XX  
 XX PA (DRIG/) DRIGGERS P H.  
 XX (RUBL/) RUBINO D M.  
 XX (SEGE/) SEGERS J.  
 XX Driggers PH, Rubino DM, Segers J;  
 DR WPI: 1999-254688/21.  
 DR P-PSDB; AAW93941.  
 XX  
 PT Breast cancer gene encoding a nuclear receptor-binding auxiliary  
 PT protein, brx  
 XX  
 PS Claim 5; Page 63-64; 69pp; English.  
 XX  
 CC This invention describes a novel human breast cancer gene encoding a  
 CC nuclear receptor-binding auxiliary protein, brx. The brx encoded protein  
 CC has anti-cancer, anti-proliferative and mitogenic activity and acts as a  
 CC transcription factor and tumour suppressor. Levels of expression of brx  
 CC can be detected using the primers, probes and antibodies (contained in  
 CC the kit) to determine a predisposition to breast and ovarian cancer and  
 CC other proliferative disorders of immune tissues. Vectors containing the  
 CC brx gene can be used for prevention and treatment of cancers and  
 CC proliferative diseases of mammalian reproductive and immune tissues.  
 XX  
 SQ Sequence 5276 BP; 1564 A; 1182 C; 1361 G; 1169 T; 0 other;

alignment\_scores:

Quality: 68.00 Length: 106  
 Ratio: 1.172 Gaps: 6  
 Percent Similarity: 54.717 Percent Identity: 29.245

alignment\_block:

US-09-528-682-4 x AAX24044 ..

Align seg 1/1 to: AAX24044 from: 1 to: 5276

23 PheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArg.....TyrSe 37  
 |||||:|||||:|||||  
 351 TTCTCTGGATGTTGGAGCTG.....GGCAGAGAGTGTCACCTC 385  
 37 rAsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG 54  
 :||:|||||:||||| ||| |||:|||||  
 386 AAAACAAGGTGACTTAAAGAGAATCTGGAGTGATCTTGACCTCTTC 435  
 54 lYserAlaAsnIleAspGly.....PheProGluValArgGlu 66  
 ||| :||| |||:|||||  
 436 ACTCACCCAGTGATGACATGGACAGCATCATCTTCCCAAGCCAGAGGAA 485  
 67 .....PheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79  
 :||:|||||:|||||  
 486 GAGCATTTGGCTGTGATATCACCAGGATCCAGTTTCATCCACCGATGACAC 535  
 79 rAlaSerLeuAsnLys.....GlnHisAspAlaAspPheLysLysTyrI 94  
 |||||:|||||:||||| ||| |||:|||||  
 536 GGCTTCACCTGGACCGACATTTCTTCATGGCAGTGATGTCTCTCTCC 585  
 94 lElYsLeuLeuIleAsnAsn.....AspGlyPhe 103  
 ||| |||:|||||



















OM of: US-09-528-682-4 to: EST:\* out\_format : pfs

Date: Jun 18, 2002 6:27 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2.1/USPTO.spool/US09528682/runat\_18062002\_082443\_7797/app\_query.fasta\_1.689  
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000  
-FGAPEXT=7.000 -YGAPO=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi  
-LIST=1000 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09528682@cgn1\_1\_5309  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-4

Query length: 110

Database: EST\*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 4056.470000

score\_list:

Strd	Orig	zScore	EScore	Len	Documentation
gb_est2:B1813876	+	82.00	186.62	487	I B1813876 PFEStoaa19d05.y1 Plasm
gb_est1:BE061488	+	78.50	178.34	471	I BE061488 MRO-BT0248-091299-103
gb_est2:BM169137	+	78.50	173.99	720	I BM169137 ESTP571660 PyBS Plasmid
gb_gss:CN5045RA	+	78.00	168.85	1055	I AL2275743 Tetraodon nigroviridis
gb_gss:CN480FOV	+	76.50	165.89	5.41	I AL1396677 T3 end of clone AROAAC
gb_gss:BN483045	+	76.00	165.65	5.57	I BH483045 BOHQU12TR BOHO Brassic
gb_est2:BF239368	+	74.00	168.62	3.81	I BP496463 AT10383 Sprime AT Dro
gb_est2:BF239368	+	74.00	164.03	6.86	I BF239368 O67PBP01 pb CDNA #20,
gb_est2:BF232289	+	74.00	162.46	8.39	I BF232289 EST543194 tomato callu
gb_gss:BN483045	+	74.00	161.62	9.35	I BH302927 CH230-58G18.TVB CHORI-
gb_gss:AZ293927	+	73.50	165.39	5.77	I BH486660 BOHKF67TR BOHK Brassic
gb_gss:AZ2523935	+	73.50	162.50	8.35	I AW687451 NF009603Rt1F1023 Deyel
gb_est1:AV687451	+	72.50	163.17	7.66	I AW687451 NF009603Rt1F1023 Deyel
gb_est1:AW257463	+	72.50	160.18	11.25	I AW257463 EST305600 KV2 Medicago
gb_est2:BG667231	+	72.50	158.98	13.11	I BG667231 EST508850 HOGA Medicago
gb_gss:BN483045	+	72.50	158.48	13.98	I BH486660 BOHKF67TR BOHK Brassic
gb_gss:BN552001	+	72.50	158.12	14.65	I BH552001 BOHMW73TF BOHM Brassic
gb_est2:BF032307	+	72.50	150.22	40.37	I BF032307 601452770F1 NIH.MGC
gb_est1:AV916505	+	72.00	159.51	12.26	I AV916505 AV916505 K. Sato unpub
gb_est1:AV916629	+	72.00	159.30	12.60	I AV916629 AV916629 K. Sato unpub
gb_est1:AV917534	+	72.00	158.05	14.78	I AV917534 AV917534 K. Sato unpub
gb_gss:BN483045	+	72.00	157.99	14.89	I BH457228 BOHK53TF BOHK Brassic
gb_est1:AV919718	+	72.00	157.96	14.95	I AV919718 AV919718 K. Sato unpub
gb_est2:BG239887	+	72.00	156.70	17.58	I BG239887 HVSME0022F14f Hordeum
gb_est2:BG236371	+	71.50	165.71	5.53	I BG26371 sae09406.y1 Gm-cl067 C
gb_est2:BF637831	+	71.50	158.53	13.90	I BF637831 NF043050F1F1036 Phosph
gb_est2:BF239122	+	71.50	158.31	14.30	I BF239122 O6BP061 pb CDNA #20,
gb_gss:AZ2006675	+	71.50	158.12	14.65	I BH552001 BOHMW73TF BOHM Brassic
gb_est2:BF636639	+	71.50	157.54	15.77	I BF636639 P11-35.D02.g1.A002 Path
gb_gss:CN5023J6	+	71.50	154.41	23.56	I AL219723 Tetraodon nigroviridis
gb_est1:AW377167	+	71.00	158.27	14.37	I AW377167 IL3-CF0220-091199-026
gb_gss:AZ408211	+	71.00	157.96	14.95	I AZ408211 IM0179K18F Mouse 10kb
gb_gss:BN483045	+	71.00	157.47	15.93	I BH37653 CH230-169119.TJ CHORI-
gb_gss:BN483045	+	71.00	156.38	16.11	I BH455623 BOGLO21TF BOGL Brassic
gb_gss:BN483045	+	71.00	157.85	17.24	I BH300761 CH230-109J14.TJ CHORI-
gb_gss:BN5595335	+	71.00	153.67	25.92	I BH5595335 BOGVO47TR BOGO Brassic
gb_gss:CN506259	+	71.00	152.74	29.21	I AL421891 T7 end of clone AY0AAC
gb_est1:AZ907481	+	70.50	157.97	14.93	I AL907481 RC-BT1144-190399-038 EST
gb_est2:BE581518	+	70.50	157.91	15.04	I BE581518 kq47a07.y1 TBH955TM-SRR
gb_est2:BJ172105	+	70.50	155.62	20.20	I BJ172105 full length d



gb\_est2:BE445930 68.00 152.29 30.93 479 ! BE445930 WHE1143\_D02\_H0325 Whea  
gb\_gss:BH350934 68.00 151.96 32.29 495 ! BG350934 CH230-65I122.TJ CHORI-2  
gb\_gss:BG730440 68.00 151.87 32.63 499 ! BG730440 df44e07.y1 Wellcome CR  
gb\_gss:BH348111 68.00 150.75 37.70 557 ! BH348111 CH230-80N8.TJ CHORI-23  
gb\_gss:BH358461 68.00 149.34 45.15 639 ! BH358461 CH230-34C18.TJ CHORI-2  
gb\_gss:BG389457 68.00 149.10 46.55 654 ! BG389457 602414155F1 NIH.MGC.92  
gb\_gss:BH270398 68.00 148.87 47.96 669 ! BH270398 CH230-1914.TJ CHORI-23  
gb\_est2:BE443809 68.00 148.84 48.15 671 ! BE443809 WHE1122.A11.A222S Whea  
gb\_gss:BH265777 68.00 148.44 50.71 698 ! BH265777 CH230-185N17.TJ CHORI-  
gb\_est2:AG112361 68.00 147.88 54.46 737 ! AG112361 Pan troglodytes DNA, C  
gb\_gss:AG396597 68.00 147.38 58.08 774 ! AG396597 TST354615 tomato flowe  
gb\_gss:BH579324 68.00 147.20 59.46 788 ! BH579324 BOHBK79R BOBK Brassic  
gb\_gss:BH331396 68.00 146.68 63.56 829 ! BH331396 CH230-94J18.TJ CHORI-2  
gb\_gss:BH467307 68.00 146.64 63.86 832 ! BH467307 BOHGT707F BOHG Brassic  
gb\_est2:BM033634 68.00 146.10 68.43 877 ! BM033634 kH6e07.y1 Ascariis su  
gb\_gss:BH262173 68.00 145.13 77.48 964 ! BH262173 CH230-50K5.TJ CHORI-23  
gb\_gss:BH271354 68.00 150.52 38.83 505 ! BH271354 CH230-19J4.TV CHORI-23  
gb\_est2:BA747490 68.00 150.01 41.47 531 ! BA747490 da142g05.y1 NICHD XGC  
gb\_gss:BA834206 68.00 149.87 42.19 538 ! BA834206 of15e05.sl NCI.CCAP.K1  
gb\_gss:BM289660 68.00 149.35 45.10 566 ! BM289660 952026001.y1 952 - BM  
gb\_gss:BH364836 68.00 148.79 48.47 598 ! BH364836 CH230-127L10.TV CHORI-  
gb\_gss:BM033634 68.00 148.62 49.54 608 ! BM033634 AUJCA49TF AUJC Arabid  
gb\_gss:BM033634 68.00 148.11 52.88 639 ! BM033634 EST503153 Plasmodium y  
gb\_gss:BH355439 68.00 146.33 66.41 760 ! BH355439 CH230-163H5.TJ CHORI-2  
gb\_est2:BG255309 68.00 143.99 89.76 956 ! BG255309 602369892F1 NIH.MGC.91  
gb\_est2:BH350301 68.00 150.72 37.83 439 ! BH350301 yv08009.y1 Soares fetal  
gb\_est2:BH370069 68.00 150.61 38.40 444 ! BH370069 EBrc08\_SQ002.006.R IGR  
gb\_est2:TA4740 68.00 149.98 41.61 472 ! TA4740 8003 Lambda-PRL2 Arabid  
gb\_est2:BM863771 68.00 149.85 42.31 478 ! BM863771 kx48b10.y1 Parastromy  
gb\_gss:BH303430 68.00 149.64 43.47 488 ! BH303430 CH230-10B9.TV CHORI-23  
gb\_gss:BH344853 68.00 149.33 45.24 503 ! BH344853 CH230-1103.TV CHORI-23  
gb\_gss:BH523029 68.00 148.81 48.33 529 ! BH523029 BOGO33TF BOGO Brassic  
gb\_est2:BE603252 68.00 148.68 49.17 536 ! BE603252 EST413555 tomato devel  
gb\_gss:BE458215 68.00 148.27 51.84 558 ! BE458215 EST413555 tomato devel  
gb\_gss:BH296814 68.00 148.10 52.94 567 ! BH296814 CH230-45C24.TJ CHORI-2  
gb\_gss:AZ416334 68.00 147.94 54.05 576 ! AZ416334 LM0191102R Mouse 10kb  
gb\_gss:BH380660 68.00 147.73 55.53 588 ! BH380660 CH230-89D19.TJ CHORI-2  
gb\_gss:BH355977 68.00 147.25 59.03 616 ! BH355977 CH230-19918.TV CHORI-2  
gb\_gss:BH317194 68.00 147.24 59.15 617 ! BH317194 CH230-182G18.TJB CHORI-  
gb\_est2:AV921915 68.00 146.89 61.81 638 ! AV921915 AV921915 K. Sato unpub  
gb\_gss:BE603758 68.00 146.88 61.94 639 ! BE603758 GS154 GS Lambda-Trip  
gb\_gss:BH318771 68.00 146.86 62.07 640 ! BH318771 CH230-20C9.TV CHORI-23  
gb\_est2:BF072383 68.00 146.78 62.70 645 ! BF072383 db56e01.x1 Blackshear/  
gb\_est2:AV920534 68.00 146.70 63.34 650 ! AV920534 AV920534 K. Sato unpub  
gb\_gss:BH319011 68.00 146.61 64.11 656 ! BH319011 AV931901 K. Sato unpub  
gb\_gss:BH355074 68.00 146.56 64.50 659 ! BH355074 CH230-199G6.TV CHORI-2  
gb\_gss:BH288066 68.00 146.50 65.01 663 ! BH288066 CH230-17116.TJ CHORI-2  
gb\_est2:AV918031 68.00 146.36 66.17 672 ! AV918031 AV918031 K. Sato unpub  
gb\_gss:BM162462 68.00 146.35 66.30 673 ! BM162462 EST564985 PyBS Plasmod  
gb\_gss:BH334945 68.00 146.11 68.38 689 ! BH334945 CH230-99F1.TJ CHORI-23  
gb\_gss:AZ991167 68.00 145.31 75.77 745 ! AZ991167 2M027510F Mouse 10kb  
gb\_gss:BM032044 68.00 145.29 75.90 746 ! BM032044 HVSMen0023009F Hordeu  
gb\_gss:BH354151 68.00 144.44 84.70 811 ! BH354151 CH230-161D19.TJ CHORI-2  
gb\_gss:BH266175 68.00 144.29 86.35 823 ! BH266175 CH230-55L8.TJ CHORI-23  
gb\_gss:BH260179 68.00 144.15 87.87 834 ! BH260179 BOHAZ92TR BOHA Brassic  
gb\_gss:BH16689 68.00 143.84 91.49 860 ! BH16689 BOHAZ92TR BOHA Brassic  
gb\_est2:BG251263 68.00 142.29 111.52 1000 ! BG251263 602365889F1 NIH.MGC.9  
gb\_hic:AK016559 68.00 132.14 410.17 2696 ! AK016559 Mus musculus adult ma  
gb\_est2:AW129870 68.00 149.70 43.12 430 ! AW129870 707003306.x2 707 - Mix  
gb\_est2:BF490859 68.00 148.81 48.33 469 ! BF490859 AT27205.5prime AT Dros  
gb\_est2:BM600148 68.00 148.53 50.09 482 ! BM600148 P11\_79\_H07.b1.A002 Pat  
gb\_est2:BF230449 68.00 148.51 50.23 483 ! BF230449 252097 BARC 5BOV Bos t  
gb\_est2:BG601077 68.00 148.32 51.46 492 ! BG601077 EST500167 Plasmodium y  
gb\_est2:BM487232 68.00 147.93 54.09 511 ! BM487232 175875 BARC 5BOV Bos t  
gb\_est2:BM428976 68.00 147.14 59.86 552 ! BM428976 952026001.y1 952 - BM  
gb\_gss:AZ524660 68.00 147.03 60.71 558 ! AZ524660 234PBH09 PB MBN #21 Pl  
gb\_gss:BH355053 68.00 146.76 62.87 573 ! BH355053 CH230-81P18.TJ CHORI-2  
gb\_est2:BF727653 68.00 146.76 63.15 575 ! BF727653 10000551A07.x3 1000 - U

1000 ! AZ219309 Sheared DNA-82H10.T  
1000 ! BE845642 2303072 BARC 5BOV Bo  
1000 ! AZ963442 2M0232008R Mouse 10  
1000 ! AZ077747 RPLC1-23-410C17.TJ R  
1000 ! BM074149 WET82-A03.T3 ISUM4  
1000 ! AW621114 707003306.x3 707 -  
1000 ! BG604058 EST503147 Plasmodi  
1000 ! BM074072 CH230-1A19.TV CHORI  
1000 ! BM078620 MEST122-B09.T3 ISUM  
1000 ! BG239556 sab69d01.y1 Unknown  
1000 ! BH310516 CH230-101F17.TV CHO  
1000 ! BH3079032 BJ079032 NIBB Mochi  
1000 ! BG505022 CH230-98B4.TJ CHORI  
1000 ! BM166487 EST569010 PyBS Plas  
1000 ! BM169297 EST569450 PyBS Plas  
1000 ! BF034629 601455060F1 NIH.MGC  
1000 ! BF052222 ENTKF22TR Entamoeba  
1000 ! AL054678 Drosophila melanoga  
1000 ! AL435277 T7 end of clone BB0  
1000 ! AL399799 T3 end of clone AS  
1000 ! AW761406 sl67b01.y1 Gm-cl027  
1000 ! AW94678 LD23553.3prime LD D  
1000 ! BG274115 WHE2231.G01.N0123 A  
1000 ! AV551467 AV551467 Arabidopsi  
1000 ! BH308035 CH230-193A1.TV CHOR  
1000 ! BH296205 CH230-187D16.TV CHOR  
1000 ! BH296566 CH230-92N14.TJB CHO  
1000 ! BH295038 CH230-104A17.TV CHO  
1000 ! BH290711 CH230-30D11.TJ CHOR  
1000 ! BH260401 CH230-29J18.TJ CHOR  
1000 ! BH342019 CH230-195C24.TJ CHO  
1000 ! BH295686 CH230-104D4.TV CHOR  
1000 ! BH777239 NX51.066.G12.F NX51  
1000 ! AW787565 94501A10.X2.945 -  
1000 ! BH337370 CH230-130B24.TV CHO  
1000 ! BH305008 CH230-25L16.TV CHO  
1000 ! BH277643 CH230-38F12.TJ CHOR  
1000 ! BH277643 CH230-38F12.TJ CHOR  
1000 ! BH27852 CH230-45022.TV CHOR  
1000 ! BH278074 CH230-48P11.TV CHOR  
1000 ! AO643064 RPLC193-ECORI-6H20.T  
1000 ! BH285105 CH230-157H10.TV CHO  
1000 ! AZ799522 2M0057K06F Mouse 10  
1000 ! AN621460 EST132258 tomato ro  
1000 ! BH309161 CH230-4L22.TV CHORI  
1000 ! BH494236 AT02677.5prime AT D  
1000 ! BH296719 CH230-92013.TJ CHOR  
1000 ! BH290903 CH230-91K21.TV CHOR  
1000 ! BH321921 CH230-7E15.TJB CHOR  
1000 ! BH294768 CH230-92G14.TJB CHO  
1000 ! BH289868 CH230-77K19.TJ CHOR  
1000 ! BH363759 CH230-164N22.TV CHO  
1000 ! BH489416 BOGPB95TF BOGP bras  
1000 ! BH547477 BOGUS18TR BOGU bras  
1000 ! BH473157 BOGUS18TR BOGU bras  
1000 ! AL197222 T7 end of clone XAX  
1000 ! AL428150 clone BA0A024A09.o  
1000 ! BH358229 CH230-175P12.TV CHO  
1000 ! BG21745 14-11 Stevia field  
1000 ! RW544487 BP330015B30H12 Soar  
1000 ! BM167760 EST570283 PyBS Plas  
1000 ! AL648120 AL648120 XGC-gastru  
1000 ! BF728127 10000556D10.x2 1000  
1000 ! BF321961 GM210008A10A2 Gm-r1  
1000 ! C77087 C77087 Mouse 3.5-dpc  
1000 ! BF423106 HC.d11.22B11.SKPL H  
1000 ! BF451147 NF098F10DRF1089 Dr



gb_est2:BM181884	65.50	143.15	99.91	641	BM181884	fv51b11.y1	Sugano STD	gb_gss:CNS071QZ	65.00	137.15	215.64	1021	AL425265	clone	BA0AB004E10
gb_gss:BH117177	65.50	142.91	103.00	656	BH117177	RPCI-24-358M13	TJ RPCI	gb_gss:A0839849	65.00	118.57	2,3e+03	6267	A0839849	69115-C76	R1KB Hom
gb_gss:BE282304	65.50	142.91	103.00	656	BE282304	CH230-108L24	TJ CHORI	gb_est1:BB282304	64.50	148.98	47.28	285	BB282304	BB282304	CH230-108L24
gb_est2:BE280140	65.50	142.82	104.23	662	BE280140	60115328F1	NIH_MGC_21	gb_gss:BA2888322	64.50	145.39	75.00	405	BA2888322	RPCI-24-158A5	TV RP
gb_gss:BH307242	65.50	142.79	104.65	664	BH307242	CH230-86A16	TJ CHORI-2	gb_gss:BA2888322	64.50	144.63	82.62	436	BA2888322	CH230-170K13	TJ CHO
gb_est1:AL652542	65.50	142.71	105.68	669	AL652542	AL652542	XGC-gastrula	gb_est1:AI489472	64.50	143.54	95.03	485	AI489472	EST247811	tomato ov
gb_est2:BG456792	65.50	142.65	106.51	673	BG456792	NF027808PL1F1063	Phosp	gb_est2:BG802056	64.50	143.29	98.13	497	BG802056	O134-09	Mouse E14.5
gb_est1:BI262933	65.50	142.64	106.72	674	BI262933	NF026610PL1F1084	Phosp	gb_est1:BI262933	64.50	143.27	98.38	498	BI262933	MGA0762f	Mga Librar
gb_est2:BM161045	65.50	142.53	108.18	681	BM161045	BOCTD34TR	BOGT Brassic	gb_est1:AI489668	64.50	143.23	98.90	500	AI489668	EST248007	tomato ov
gb_gss:BM575539	65.50	142.45	109.22	686	BM575539	ESG563568	PyBS Plasmoc	gb_gss:BM575539	64.50	143.01	101.77	511	BM575539	CH230-651T0	TJ CHOR
gb_gss:BM421472	65.50	142.18	113.21	705	BM421472	BOGK65TF	BOGK Brassic	gb_est1:AW065386	64.50	142.37	110.49	544	AW065386	614055B07.y1	614 -
gb_gss:BM597761	65.50	141.42	124.74	759	BM597761	BOGK52TR	BOGK Brassic	gb_est1:BE235178	64.50	142.37	110.49	544	BE235178	142548	MARC 1P1G SU
gb_gss:BM461874	65.50	141.23	127.77	759	BM461874	BOGK52TR	BOGK Brassic	gb_gss:BM597761	64.50	142.37	110.49	544	BM597761	CH230-123B21	TV CHO
gb_gss:BM4311026	65.50	140.47	140.94	833	BM4311026	BOGK68TF	BOGK Brassic	gb_est2:BM340471	64.50	142.09	114.51	559	BM340471	SnESTaa47c02.y1	CSN
gb_est2:BI404318	65.50	140.12	147.42	862	BI404318	CH230-101H13	TV CHORI-2	gb_gss:AY564243	64.50	141.80	118.83	575	AY564243	RPCI-23-207J10	TJ CSN
gb_est1:BE126489	65.00	150.77	37.60	270	BE126489	MI-P-CP1-nwy-a-04-0-U		gb_gss:BM340669	64.50	141.73	119.91	579	BM340669	CH230-16B17	TVB CHO
gb_est1:AW910159	65.00	148.97	47.38	322	AW910159	DEPA2037	Rat Lambda MA	gb_est2:BJ181042	64.50	141.62	121.55	585	BJ181042	BJ181042	unpublishe
gb_est1:AW910159	65.00	145.56	73.32	449	AW910159	ur75505.y1	NCI_CGAP	gb_est2:BJ181042	64.50	141.53	122.92	590	BJ181042	BJ181042	normalized
gb_est1:BE061494	65.00	145.03	78.51	473	BE061494	MRO-B70248	-091295-103-	gb_est1:AW031562	64.50	141.50	123.46	592	AW031562	AU031562	Rice cDNA
gb_gss:BM287098	65.00	144.90	79.82	479	BM287098	CH230-62B6	TJ CHORI-23	gb_est1:AW031562	64.50	141.13	129.52	614	AW031562	BJ173006	full lengt
gb_gss:BM330284	65.00	144.61	82.90	493	BM330284	CH230-201C24	TV CHORI	gb_est2:BM334811	64.50	141.13	129.52	614	BM334811	MEST142-E01	TJ ISUM
gb_est2:BJ068605	65.00	143.79	92.07	534	BJ068605	NIIBB	Mochli n	gb_gss:BM303573	64.50	140.98	132.02	623	BM303573	CH230-25H10	TV CHOR
gb_gss:BM338923	65.00	143.58	94.57	545	BM338923	CH230-51G10	TJ CHORI-2	gb_est1:AW686607	64.50	140.57	139.02	648	AW686607	NF039H12NRF1000	NO
gb_gss:BM304311	65.00	143.37	97.08	556	BM304311	CH230-14A21	TV CHORI-2	gb_gss:BM304311	64.50	140.54	139.58	650	BM304311	Pan troglodytes	DNA
gb_est2:BI617109	65.00	143.23	98.92	564	BI617109	RH47342	Sprline RH Dros	gb_gss:BM294576	64.50	140.39	142.41	660	BM294576	CH230-104A12	TJ CHO
gb_gss:BM286321	65.00	143.16	99.84	568	BM286321	CH230-62A16	TJ CHORI-2	gb_gss:BM166408	64.50	140.31	143.83	665	BM166408	BJ166408	full lengt
gb_gss:BM304311	65.00	142.98	102.15	578	BM304311	CH230-200P21	TJ CHORI	gb_gss:BM294441	64.50	140.28	144.40	667	BM294441	CH230-173A21	TJ CHO
gb_gss:BM3325290	65.00	142.91	103.08	582	BM3325290	CH230-93E24	TV CHORI-2	gb_est2:BM239751	64.50	140.13	147.25	677	BM239751	K0539F06-3	NTA Mous
gb_gss:BM358248	65.00	142.89	103.32	583	BM358248	CH230-175P18	TJ CHORI	gb_gss:BM308193	64.50	139.95	150.68	688	BM308193	CH230-4K22	TV CHORI
gb_est1:AI861114	65.00	142.89	103.32	583	AI861114	603012C10.x1	603 - str	gb_est2:C93316	64.50	139.67	156.16	708	C93316	C93316	Diclyostelium
gb_est1:AW643308	65.00	142.80	104.48	586	AW643308	CH230-23C22	W1 Blackshear	gb_gss:AY332164	64.50	139.65	165.47	709	AY332164	IM0060F01R	Mouse 1D
gb_gss:BM361781	65.00	142.66	106.35	596	BM361781	CH230-23C22	TJ CHORI-2	gb_est2:BF489221	64.50	139.37	162.27	729	BF489221	AT25155	Sprline AT D
gb_gss:AY780495	65.00	142.63	106.82	598	AY780495	2M0017119R	Mouse 10kb	gb_est2:BF489221	64.50	139.26	164.62	737	BF489221	BJ166431	full lengt
gb_gss:BM337121	65.00	142.49	108.70	606	BM337121	CH230-111C5	TJ CHORI-2	gb_gss:BM335868	64.50	139.24	164.91	738	BM335868	CH230-95E15	TJ ISUM
gb_est2:BI617016	65.00	142.44	109.41	609	BI617016	RH47217	Sprline RH Dros	gb_est2:BM348055	64.50	139.21	165.50	740	BM348055	MEST285-H11	T3 ISUM
gb_est2:BI637042	65.00	142.41	109.88	611	BI637042	SD18930	Sprline SD Dros	gb_gss:BM269802	64.50	139.21	165.50	740	BM269802	CH230-6L20	TV CHORI
gb_est2:BI637917	65.00	142.41	109.88	611	BI637917	SD20073	Sprline SD Dros	gb_gss:BM269802	64.50	139.21	165.50	740	BM269802	CH230-6L20	TV CHORI
gb_gss:BM287274	65.00	142.29	111.54	618	BM287274	CH230-12B34	TV CHORI-2	gb_gss:BM232880	64.50	139.20	165.79	741	BM232880	CH230-5318	TV CHORI
gb_est1:AI543658	65.00	142.28	111.77	619	AI543658	SD10528	Sprline SD Dros	gb_est2:BF238504	64.50	138.31	185.75	808	BF238504	601904821F1	NIH_MGC
gb_est1:BE189710	65.00	142.23	112.49	622	BE189710	db59a10.y1	Wellcome CR	gb_gss:BM287192	64.50	138.19	188.77	818	BM287192	CH230-62C18	TV CHOR
gb_gss:BM353564	65.00	142.14	113.67	627	BM353564	CH230-175B19	TJ CHORI	gb_gss:CNS000173	64.50	136.70	228.48	945	BM076342	Drosophila melanoga	
gb_gss:BM309647	65.00	142.00	115.82	636	BM309647	CH230-11B124	TJB CHORI	gb_gss:CNS02R13	64.50	136.64	230.39	952	AL210612	Tetradodon nigrovir	
gb_est2:BI229292	65.00	141.95	116.54	639	BI229292	RE27048	Sprline RE Dros	gb_est2:AY453384	64.50	136.43	236.44	971	BG445384	GA_Ea0027P03F	Goss
gb_est1:BI153021	65.00	141.82	118.46	647	BI153021	60291821F1	NCI_CGAP	gb_gss:CNS0775A1	64.50	136.42	236.76	972	AL432449	T3 end of clone BB0	
gb_gss:BM261337	65.00	141.82	118.46	647	BM261337	CH230-113G17	TJ CHORI	gb_est2:BF295310	64.50	135.93	252.23	1020	BF295310	602256289F1	NIH_MGC
gb_est2:BI623823	65.00	141.81	118.70	648	BI623823	RH56172	Sprline RH Dros	gb_gss:CNS06PYZ	64.50	135.39	270.24	1075	AL410001	T3 end of clone AW	
gb_est2:BG640644	65.00	141.76	119.42	651	BG640644	SD11774	Sprline SD Dros	gb_gss:CNS07CDM	64.50	135.33	272.23	1081	AL439040	T3 end of clone BD	
gb_gss:BM574477	65.00	141.70	120.39	655	BM574477	BOGRW63TF	BOGR Brassic	gb_gss:CNS00M6W	64.50	135.33	272.56	1082	AL079212	Drosophila melanog	
gb_est1:BB198674	65.00	141.27	127.19	683	BB198674	BB198674	R1KEN full-le	gb_est2:BI316585	64.00	146.44	65.53	324	BI316585	safo5c02.y1	Gm-cl05
gb_est2:BJ069651	65.00	141.27	127.19	683	BJ069651	BJ069651	NIIBB Mochli n	gb_est2:BI094918	64.00	146.07	68.73	336	BI094918	sa601901.y1	Gm-cl05
gb_gss:BM261460	65.00	141.10	129.89	694	BM261460	CH230-9014	TJ CHORI-23	gb_est1:AV554527	64.00	145.92	70.08	341	BM169807	EST572330	PyBS Plas
gb_gss:BM326758	65.00	141.09	129.89	694	BM326758	CH230-32D15	TJ CHORI-2	gb_gss:AY4447865	64.00	145.36	75.25	360	AV554527	AV554537	Arabidops1
gb_gss:BM3531100	65.00	141.05	130.13	695	BM3531100	CH230-54M18	TJ CHORI-2	gb_gss:BI700106	64.00	145.30	75.80	362	AZ447865	IM0245A15F	Mouse 10
gb_gss:BM327343	65.00	140.83	134.58	713	BM327343	CH230-90D23	TJ CHORI	gb_est1:BI770106	64.00	145.28	76.07	363	BI7700106	sa603c02.y1	Gm-cl08
gb_gss:BM260072	65.00	140.66	137.56	725	BM260072	CH230-90D23	TJ CHORI-2	gb_est2:BI773366	64.00	144.86	80.23	378	AI777366	EST263774	tomato se
gb_gss:BM295795	65.00	140.31	143.82	750	BM295795	CH230-9204	TJB CHORI-2	gb_est1:AA497735	64.00	144.33	85.85	398	BI788118	sa699h04.y1	Gm-cl06
gb_gss:BM270652	65.00	140.27	145.58	753	BM270652	CH230-83AT	TV CHORI-23	gb_est1:AI443585	64.00	143.93	90.41	414	AT443585	sa42c02.y1	Gm-cl004
gb_gss:BM303478	65.00	140.21	145.58	757	BM303478	CH230-72D15	TV CHORI-2	gb_est1:AI737029	64.00	143.81	91.84	419	AT496659	sb09g03.y1	Gm-cl004
gb_gss:BM439369	65.00	139.84	146.34	760	BM439369	CH230-120I12	TJ CHORI	gb_est1:BI737029	64.00	143.73	92.71	422	BI737029	sb36g10.y1	Gm-cl013
gb_gss:BM322725	65.00	139.84	146.34	760	BM322725	CH230-76G17	TJB CHORI-2	gb_est2:BF803636	64.00	143.54	95.02	440	BF598648	sv20a12.y1	Gm-cl057
gb_gss:BM273098	65.00	139.76	154.23	791	BM273098	CH230-83H10	TJ CHORI-2	gb_est2:BF803636	64.00	143.17	99.69	446	BF803636	QV1-C10175-021100-4	
gb_gss:BM506X08	65.00	139.64	156.80	801	BM506X08	T7 end of clone AX0A0		gb_est2:BF803636	64.00	143.10	100.57	449	BF811971	QV1-C10175-021100-5	
gb_gss:BM358496	65.00	139.59	157.83	805	BM358496	CH230-176B13	TJ CHORI	gb_est2:BF803636	64.00	143.03	101.46	452	BF803464	QV1-C10175-021100-4	
gb_gss:BM299933	65.00	139.43	160.92	817	BM299933	CH230-179B13	TJ CHORI	gb_est2:BG134452	64.00	142.65	106.50	469	BI54452	sa699d06.y1	Gm-cl07
gb_gss:BM311521	65.00	139.40	161.70	820	BM311521	CH230-62K18	TJ CHORI-2	gb_gss:BM280392	64.00	142.65	106.50	469	BI280392	CH230-191H10	TV CHO
gb_gss:BM277652	65.00	139.31	163.51	827	BM277652	CH230-176G17	TJ CHORI	gb_est2:BG507944	64.00	142.59	107.39	472	BG507944	sa683c04.y1	Gm-cl07
gb_gss:BM2167359	65.00	139.14	167.												



gb_gss: BH340242	64.00	142.06	114.92	497	! BH340242	CH230-41E22.TJ	CHORI-2	gb_gss: BH262780	64.00	137.26	212.61	794	! BH262780	CH230-171E21.TJ	CHOR
gb_est2: BG726797	64.00	142.02	115.53	499	! BG726797	sac04d12.y1	Gm-cl0107.G	gb_gss: BH264097	64.00	137.24	213.32	796	! BH264097	CH230-91C13.TJ	CHOR
gb_est1: A900348	64.00	141.71	120.11	514	! A900348	sc04d12.y1	Gm-cl0102.G	gb_gss: A0935402	64.00	137.06	218.26	810	! A0935402	CPG24542B	CPiOWAGDNA
gb_gss: BH531381	64.00	141.67	120.11	514	! BH531381	CH230-54N23.TJ	CHORI-2	gb_gss: BH337749	64.00	136.94	221.45	819	! BH337749	CH230-51C2.TJ	CHORI
gb_gss: A9228051	64.00	141.67	120.73	516	! A9228051	479.dif03h12.s1	Saccha	gb_gss: BH570197	64.00	136.92	221.45	821	! BH570197	BOGYO11TR	BOGY Bras
gb_gss: AQ502574	64.00	141.46	124.12	527	! AQ502574	V5H2 mtm-3xHA/lacZ	Inf	gb_gss: BH431432	64.00	136.89	222.87	823	! BH431432	CH230-52F20.TV	CHOR
gb_est2: B162110	64.00	141.28	126.91	536	! B162110	RE01521.Sprime	RE Dros	gb_gss: BH073736	64.00	136.47	235.40	858	! BH073736	RPCI-24-391E18.TJ	R
gb_est2: B1972255	64.00	141.17	128.78	542	! B1972255	sag89b10.y1	Gm-cl0184.C	gb_gss: BH357701	64.00	136.38	237.92	865	! BH357701	CH230-6B34.TJ	CHORI
gb_est1: A1259492	64.00	141.11	129.71	545	! A1259492	LP02916.Sprime	LP Dros	gb_gss: BH072684	64.00	136.36	238.64	867	! BH072684	RPCI-24-391C18.TJ	R
gb_est2: BM188049	64.00	141.11	129.71	545	! BM188049	sap84b06.y1	Gm-cl0174.C	gb_gss: BH339740	64.00	136.12	246.26	868	! BH339740	CH230-130I19.TV	CHOR
gb_gss: BH333605	64.00	141.10	130.03	546	! BH333605	CH230-26L5.TV	CHORI-2	gb_gss: BH3339740	64.00	135.56	264.63	938	! BH3339740	T7 end of clone	BCO
gb_est1: A9309043	64.00	141.06	130.65	548	! A9309043	sf94b08.y1	Gm-cl0109.G	gb_gss: C9NS0784	64.00	133.24	286.32	938	! C9NS0784	Mus musculus	0 day
gb_est2: B1587261	64.00	141.06	130.65	548	! B1587261	Sprime RH	Dros	gb_hct: AK020784	64.00	134.94	286.32	938	! AK020784	T3 end of clone	XAR
gb_gss: BH357704	64.00	141.01	131.59	551	! BH357704	CH230-163M18.TV	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH3505192	64.00	140.97	132.22	553	! BH3505192	CH230-192X21.TJ	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: BM578296	64.00	140.95	132.53	554	! BM578296	rk10c10.y1	Meloidogvne	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1235763	64.00	140.82	134.74	561	! B1235763	RE31684.Sprime	RE Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH297578	64.00	140.82	134.74	561	! BH297578	CH230-1K11.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1612980	64.00	140.73	136.32	566	! B1612980	RH42019.Sprime	RH Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH338054	64.00	140.67	137.27	569	! BH338054	CH230-40P22.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH366654	64.00	140.66	137.58	570	! BH366654	CH230-112H19.TJ	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est1: A9348676	64.00	140.62	138.22	572	! A9348676	GM210003A11E10	Gm-cl0102	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH348501	64.00	140.57	139.17	575	! BH348501	CH230-42G18.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1700119	64.00	140.48	140.76	580	! B1700119	sag63d04.y1	Gm-cl0182.C	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1757125	64.00	140.48	140.76	580	! B1757125	GM027688F1	NIH.MGC.62	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH310410	64.00	140.48	140.76	580	! BH310410	CH230-26FA.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1815289	64.00	140.42	141.72	583	! B1815289	pfEST04a16g07.y1	Plasm	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1481789	64.00	140.37	142.68	586	! B1481789	RE64390.Sprime	RE Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH233164	64.00	140.35	143.00	587	! BH233164	CH230-116B12.TJ	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1797798	64.00	140.34	143.32	588	! B1797798	CH230-84D3.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH221728	64.00	140.30	143.96	590	! BH221728	f185a10.y1	Sugano kaw	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH328115	64.00	140.20	145.88	596	! BH328115	CH230-124I118.TVB	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est1: A9035694	64.00	140.18	146.20	597	! A9035694	ESW281848	tomato callu	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH346257	64.00	140.17	147.81	602	! BH346257	CH230-8L5.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH223205	64.00	140.08	148.14	603	! BH223205	CH230-116D15.TJB	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH253380	64.00	139.98	150.08	609	! BH253380	CH230-197M2.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: A9647646	64.00	139.88	152.02	615	! A9647646	EST307124	tomato germi	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH289042	64.00	139.88	152.02	615	! BH289042	CH230-2M17.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est1: A9331798	64.00	139.86	152.35	616	! A9331798	614097C01.x1	614 - roc	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH312118	64.00	139.83	152.99	618	! BH312118	CH230-66H18.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH3597375	64.00	139.76	154.30	622	! BH3597375	ESW496053	CSNS Solanum	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1231101	64.00	139.68	155.93	627	! B1231101	RE17989.Sprime	RE Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1570335	64.00	139.66	156.25	628	! B1570335	RH02643.Sprime	RH Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1588651	64.00	139.66	156.25	628	! B1588651	RH03039.Sprime	RH Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1521573	64.00	139.63	156.58	629	! B1521573	RE21511.Sprime	RE Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1587799	64.00	139.63	156.91	630	! B1587799	RE28675.Sprime	RE Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est1: A9820464	64.00	139.53	158.87	636	! A9820464	LD24208.Sprime	LD Dros	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH319406	64.00	139.53	158.87	636	! BH319406	CH230-79H8.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B1527441	64.00	139.29	163.81	651	! B1527441	BOHC059TF	BOHC Brassic	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH343720	64.00	139.23	165.13	655	! BH343720	CH230-151I22.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_est2: B178432	64.00	139.22	165.47	656	! B178432	CH230-191D19.TV	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH278951	64.00	139.12	167.46	662	! BH278951	CH230-17J7.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH266292	64.00	139.05	169.12	667	! BH266292	CH230-55M7.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH268459	64.00	138.58	179.51	698	! BH268459	CH230-82M24.TJ	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH364683	64.00	138.48	181.88	705	! BH364683	CH230-60E16.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH360243	64.00	138.32	182.56	707	! BH360243	CH230-104H16.TV	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH348990	64.00	138.31	185.62	716	! BH348990	CH230-18K13.TVB	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH3434986	64.00	138.23	187.66	722	! BH3434986	CH230-150I11.TV	CHORI-2	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: BH278959	64.00	138.22	188.00	723	! BH278959	CH230-173F14.TJ	CHORI-	gb_est1: A9317574	64.00	144.36	85.57	352	! A9317574	sg52h01.y1	Gm-cl0136
gb_gss: AG015249	64.00	138.14	190.06	729	! AG015249	Homo sapiens	genomic	gb_gss: AG015249	64.00	138.14	190.06	729	! AG015249	Homo sapiens	genomic
gb_gss: BH322113	64.00	137.96	194.52	742	! BH322113	CH230-159P7.TV	CHORI-2	gb_gss: BH270960	64.00	137.84	197.29	665	! BH270960	CH230-162K5.TJ	CHOR
gb_gss: BH3553382	64.00	137.93	195.21	744	! BH3553382	BOHBS11TR	BOHB Brassic	gb_gss: A0034215	64.00	137.62	203.16	680	! A0034215	CPG0508B	CPiOWAGDNA
gb_gss: BH328008	64.00	137.91	195.55	745	! BH328008	CH230-45P20.TV	CHORI-2	gb_est1: A9317574	64.00	137.41	208.67	694	! A9317574	A0060684	Diclyostel
gb_gss: BH317206	64.00	137.90	195.90	746	! BH317206	CH230-119K14.TV	CHORI-	gb_est1: A9317574	64.00	137.41	208.67				







[illegible]



gb_gss:AZ793987	62.00	134.51	302.64	642	1	AZ793987	2M0047J04R	Mouse	10kb	gb_hlc:AK013042	62.00	125.61	947.45	1531	1	AK013042	Mus musculus	10, 1	
gb_gss:BH297846	62.00	134.38	307.62	650	1	BH297846	CH230-57111.TV	CHORI	1	gb_hlc:AK004957	62.00	118.16	2.5e+03	3170	1	AK004957	Mus musculus	adult	
gb_est2:BG456060	62.00	134.37	308.22	651	1	BG456060	NF073F07PL1F1001	Phosp	1	gb_est1:AV116742	61.50	141.00	131.66	302	1	AV116742	Mus musculus		
gb_gss:BH270249	62.00	134.35	308.84	652	1	BH270249	CH230-19115.TV	CHORI-2	2	gb_est2:BG155825	61.50	140.83	134.53	307	1	BG155825	saet77005.y1	Gm-cl06	
gb_est2:BF639420	62.00	134.29	311.33	656	1	BF639420	NF0118021N1F1014	Insec	1	gb_est2:BG2653106	61.50	139.72	155.02	342	1	BG2653106	saet82801.y1	Gm-cl05	
gb_gss:BH309516	62.00	134.26	312.58	658	1	BH309516	CH230-98K11.TJ	CHORI-2	2	gb_est1:AA332813	61.50	139.61	157.41	346	1	AA332813	EST36831	Embryo, 8	
gb_est1:AW691578	62.00	134.23	313.83	660	1	AW691578	NF046P115T1F1000	Insec	1	gb_gss:BH356893	61.50	139.23	165.22	359	1	BH356893	CH230-47C3	TJ CHORI	
gb_est1:AW641740	62.00	134.23	313.83	660	1	AW641740	NF066D07N1F1001	Devel	1	gb_est2:C60641	61.50	139.20	165.82	360	1	C60641	C60641	Yuji Kohara un	
gb_est2:BG1263436	62.00	134.21	314.45	661	1	BG1263436	NF091802PL1F1025	Phosp	1	gb_est2:Z18503	61.50	138.81	174.34	374	1	Z18503	ATTS0789	AC16H Arabid	
gb_gss:AG161329	62.00	134.21	314.45	661	1	AG161329	NF091802PL1F1025	Phosp	1	gb_est2:BM278553	61.50	138.65	178.02	380	1	BM278553	as_tgz_64B04	SKPL A	
gb_est1:AL655348	62.00	134.19	315.08	662	1	AL655348	AL655348	XGC-gastrula	1	gb_est2:BM450619	61.50	138.12	190.43	400	1	BM450619	kt28d08.y3	Strongyl	
gb_est2:BG457186	62.00	134.19	315.08	662	1	BG457186	NF100803PL1F1027	Phosp	1	gb_gss:AO863434	61.50	137.60	203.66	421	1	AO863434	nbe0202N22r	CUGI R	
gb_gss:BH1758946	62.00	134.19	315.08	662	1	BH1758946	G03042414F1	NIH_MGC_11	1	gb_est2:BI1480348	61.50	137.05	218.40	444	1	BI1480348	CH230-177G20	TJ CHORI	
gb_est2:BG456308	62.00	134.13	317.58	666	1	BG456308	NF077006PL1F1048	Phosp	1	gb_est1:BE190150	61.50	136.30	240.62	478	1	BE190150	sa08b02.y1	Gm-cl035	
gb_gss:BH1264336	62.00	134.10	318.83	668	1	BH1264336	NF11C10PL1F1082	Phosp	1	gb_gss:BH332846	61.50	136.23	242.60	481	1	BH332846	CH230-126A18	TV CHOR	
gb_gss:BH297452	62.00	134.00	323.23	675	1	BH297452	CH230-123M11	TVB CHORI	1	gb_gss:BH255921	61.50	136.13	245.92	486	1	BH255921	CH230-170L15	TJ CHOR	
gb_est2:BM266166	62.00	133.98	323.86	676	1	BM266166	NF101A11N1F1085	Insec	1	gb_gss:BH355921	61.50	136.02	249.25	493	1	BH355921	CH230-175D4	TJ CHOR	
gb_gss:BH324001	62.00	133.98	323.86	676	1	BH324001	CH230-8086	TVB CHORI-23	23	gb_est2:BI815690	61.50	135.98	250.58	493	1	BI815690	PfEST0aa31e12.y1	PI	
gb_gss:BH346084	62.00	133.94	325.74	679	1	BH346084	CH230-8LJ10	TVB CHORI-2	2	gb_gss:BH273045	61.50	135.92	252.59	496	1	BH273045	CH230-177G20	TJ CHOR	
gb_gss:BH310093	62.00	133.88	328.27	683	1	BH310093	CH230-86P22	TJ CHORI-2	2	gb_est2:BG226448	61.50	135.84	255.26	500	1	BG226448	kp99c08.y1	TBN95TM-	
gb_gss:AG172792	62.00	133.88	328.27	683	1	AG172792	Pan troglodytes DNA	1	1	gb_est2:BM169839	61.50	135.73	258.62	505	1	BM169839	EST572362	PyBS Plas	
gb_gss:BH351424	62.00	133.74	333.96	692	1	BH351424	CH230-106D24	TJ CHORI	1	gb_est1:AA981368	61.50	135.65	261.31	509	1	AA981368	6M0054A09	x1 660 -	
gb_gss:BH352686	62.00	133.73	334.59	693	1	BH352686	CH230-106G12	TJ CHORI	1	gb_est1:AA981368	61.50	135.61	262.66	511	1	AA981368	vx56c04.r1	Stratage	
gb_gss:BH322394	62.00	133.70	335.86	695	1	BH322394	CH230-132N4	TVB CHORI-2	2	gb_est2:BM278609	61.50	135.57	264.01	513	1	BM278609	As_tgz_65A04	SKPL A	
gb_gss:BH3641328	62.00	133.68	336.50	696	1	BH3641328	NF051H08N1F1075	Insec	1	gb_est2:BM196786	61.50	135.55	264.69	514	1	BM196786	BJ196786	normalized	
gb_est1:AL646639	62.00	133.61	339.67	701	1	AL646639	AL646639	XGC-neurula	9	gb_gss:BH318886	61.50	135.43	268.75	520	1	BH318886	k161e08.y1	Ascaris	
gb_gss:BH363051	62.00	133.56	341.58	704	1	BH363051	CH230-127C19	TVB CHORI	1	gb_est2:BI194884	61.50	135.41	269.43	521	1	BI194884	fullid07.y3	Campbell	
gb_gss:BH346223	62.00	133.51	344.13	708	1	BH346223	CH230-120C12	TVB CHORI	1	gb_est2:BF228444	61.50	135.41	269.43	521	1	BF228444	As_tgz_27H07	SKPL A	
gb_gss:BH312492	62.00	133.46	346.05	711	1	BH312492	CH230-3LJ3	TJ CHORI-23	23	gb_gss:BH279318	61.50	135.40	270.11	522	1	BH279318	RPC1-24	349M15	
gb_gss:BH320389	62.00	133.41	348.61	715	1	BH320389	CH230-53C6	TJ CHORI-23	23	gb_est2:BM278731	61.50	135.30	273.52	527	1	BM278731	As_tgz_66F02	SKPL A	
gb_gss:BH342999	62.00	133.35	351.17	719	1	BH342999	CH230-16L13	TVB CHORI-23	23	gb_gss:BH078556	61.50	135.09	281.04	538	1	BH078556	PC1-24	377M1	
gb_gss:BH309488	62.00	133.29	353.74	723	1	BH309488	EST530898	GPOD Medica	9	gb_gss:BH27362	61.50	134.94	286.54	546	1	BH27362	CH230-105F10	TJ CHOR	
gb_gss:BH345113	62.00	133.25	355.67	726	1	BH345113	CH230-120A12	TVB CHORI	1	gb_gss:BH350156	61.50	134.94	286.54	546	1	BH350156	CH230-32N20	TJ CHOR	
gb_gss:BH281257	62.00	133.22	356.96	728	1	BH281257	CH230-109F14	TVB CHORI	1	gb_gss:BH350156	61.50	134.88	288.61	549	1	BH350156	CH230-106C4	TJ CHOR	
gb_gss:BH312274	62.00	133.10	362.76	737	1	BH312274	CH230-10J3	TJ CHORI-2	2	gb_est1:AW927915	61.50	134.86	289.30	550	1	AW927915	945011A10.y1	945 -	
gb_gss:BH376546	62.00	133.07	362.76	737	1	BH376546	AG-ND-138H7	TF ND-TAM	1	gb_est2:BF429718	61.50	134.82	290.68	552	1	BF429718	253567	MARC BSM Bos	
gb_gss:BH353201	62.00	133.07	364.06	739	1	BH353201	CH230-221D20	TJ CHORI	1	gb_est2:BM277735	61.50	134.75	293.45	556	1	BM277735	K0248B09-3	NIA Mous	
gb_gss:BH278942	62.00	133.03	366.04	742	1	BH278942	CH230-17J3	TJ CHORI-23	23	gb_est1:AW787564	61.50	134.71	294.84	558	1	AW787564	945011A10.x1	945 -	
gb_gss:BH288984	62.00	133.01	366.64	743	1	BH288984	CH230-77E8	TJ CHORI-23	23	gb_est2:BM078784	61.50	134.71	294.84	558	1	BM078784	MEST124-C01	TJ ISUM	
gb_gss:BH253666	62.00	132.99	367.94	745	1	BH253666	602975359F1	NIH_MGC_12	12	gb_gss:BH352678	61.50	134.66	296.92	561	1	BH352678	CH230-199A7	TJ CHOR	
gb_gss:BH272658	62.00	132.96	369.24	747	1	BH272658	CH230-56K16	TVB CHORI-23	23	gb_est1:BM053204	61.50	134.64	297.61	562	1	BM053204	md6f09.x1	Melton N	
gb_gss:BH264674	62.00	132.96	369.24	747	1	BH264674	CH230-84L9	TVB CHORI-23	23	gb_est1:AA400364	61.50	134.55	301.10	567	1	AA400364	md17b02.r1	Beddingt	
gb_gss:BH265304	62.00	132.92	371.19	749	1	BH265304	CH230-189K10	TVB CHORI	1	gb_est1:AA4199297	61.50	134.55	301.10	567	1	AA4199297	AU199297	unpublishe	
gb_gss:BH265304	62.00	132.92	371.19	750	1	BH265304	CH230-199K10	TVB CHORI	1	gb_est2:BM18930	61.50	134.55	301.10	567	1	BM18930	k162b04.y1	Ascaris	
gb_est2:BE997398	62.00	132.82	375.74	757	1	BE997398	ESP429121	GVSN Medica	9	gb_gss:BH352678	61.50	134.66	296.92	561	1	BH352678	CH230-199A7	TJ CHOR	
gb_gss:BH302599	62.00	132.82	375.74	757	1	BH302599	CH230-97N18	TVB CHORI-23	23	gb_est1:BM426525	61.50	134.51	302.49	569	1	BM426525	pgf2n.pk02	p15 Nor	
gb_gss:BH260261	62.00	132.66	383.58	769	1	BH260261	CH230-13D13	TJC CHORI	1	gb_est1:AA695430	61.50	134.49	303.19	570	1	AA695430	GMO2818	5prime GM	
gb_gss:BH311481	62.00	132.61	386.21	773	1	BH311481	CH230-62K13	TJ CHORI-23	23	gb_est1:AA440019	61.50	134.44	305.29	573	1	AA440019	LD15087	5prime LD D	
gb_gss:BH297940	62.00	132.57	388.18	776	1	BH297940	CH230-84F13	TVB CHORI-23	23	gb_est1:BG082113	61.50	134.44	305.29	573	1	BG082113	H3072F04-5	NIA Mous	
gb_gss:BH308820	62.00	132.57	388.18	776	1	BH308820	CH230-38D22	TJ CHORI-23	23	gb_est2:BG262649	61.50	134.44	305.29	573	1	BG262649	WHE0938	F04_K08Z5 W	
gb_gss:BH310800	62.00	132.41	396.08	788	1	BH310800	CH230-4P13	TJ CHORI-23	23	gb_gss:BH331977	61.50	134.42	305.99	574	1	BH331977	CH230-46014	TJ CHOR	
gb_gss:BH366023	62.00	132.41	396.08	788	1	BH366023	CH230-28J14	TVB CHORI-23	23	gb_est2:BI066942	61.50	134.41	306.69	575	1	BI066942	pgfln.pk010	ea norm	
gb_gss:BH331512	62.00	132.40	396.74	789	1	BH331512	CH230-50D23	TVB CHORI-23	23	gb_gss:AO855749	61.50	134.26	312.30	583	1	AO855749	cp61563A	CpioWAGNA LD	
gb_gss:BH274124	62.00	132.32	400.70	795	1	BH274124	CH230-83J11	TVB CHORI-23	23	gb_est1:AA820971	61.50	134.25	313.01	584	1	AA820971	LD24993	5prime LD D	
gb_gss:BH315985	62.00	132.31	401.37	796	1	BH315985	CH230-88G12	TVB CHORI-23	23	gb_gss:BH313447	61.50	134.19	315.12	587	1	BH313447	CH230-193N15	TJ CHOR	
gb_gss:BH490388	62.00	132.27	403.37	799	1	BH490388	BOHB1317R	BOHB Brassic	1	gb_gss:AG665727	61.50	134.16	316.53	589	1	AG665727	IM0547N13F	Mouse 10	
gb_gss:BH26515	62.00	132.23	405.34	802	1	BH26515	BOHBV26TR	BOHB Brassic	1	gb_gss:BH301068	61.50	134.06	320.77	595	1	BH301068	CH230-36J1	TV CHORI	
gb_gss:BH325640	62.00	132.22	406.01	803	1	BH325640	CH230-106C13	TJ CHORI	1	gb_est1:AW586503	61.50	133.99	323.60	599	1	AW586503	EST318198	MHAM Medi	
gb_est1:AL545961	62.00	132.21	406.67	804	1	AL545961	AL545961	LT1_NFL1006	PI	1	gb_est1:AI964474	61.50	133.88	327.87	605	1	AI964474	496011C08.x1	496 -
gb_gss:BH362855																			



gb\_est2:BM282772 + 61.50 133.01 366.82 659 ! BM282772 k140d08.y1 Ascaris suu  
 gb\_gss:BH319636 + 61.50 132.96 369.02 662 ! BH319636 CH230-159K12.TJB CHORI  
 gb\_est1:AV672550 + 61.50 132.89 372.68 667 ! AV672550 AV672550 Nori Satoh un  
 gb\_gss:BH319724 + 61.50 132.87 373.41 668 ! BH319724 CH230-20E23.TJ CHORI-2  
 gb\_gss:BH467561 - 61.50 132.87 373.41 668 ! BH467561 BOGP661TR BOGP Brassic  
 gb\_gss:AW773956 + 61.50 132.82 375.62 671 ! AW773956 EST332942 KV3 Medicag  
 gb\_est2:BG918770 + 61.50 132.81 376.35 672 ! BG918770 602819249F1 NCI\_CGAP\_M  
 gb\_est2:C84701 + 61.50 132.79 377.09 673 ! C84701 C84701 Dictyostelium dis  
 gb\_est2:BM160882 + 61.50 132.75 379.30 676 ! BM160882 EST563405 PYBS Plasmod  
 gb\_est1:BB638635 - 61.50 132.72 380.77 678 ! BB638635 BB638635 RIKEN full-le  
 gb\_gss:BH318304 + 61.50 132.69 382.25 680 ! BH318304 CH230-79F18.TJ CHORI-2

seq\_name: gb\_est2:BI813876

seq\_documentation\_block:

LOCUS BI813876 487 bp mRNA linear EST 03-OCT-2001  
 DEFINITION PfESToaal9d05.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium

ACCESSION BI813876

VERSION BI813876.1 GI:15906057

KEYWORDS EST.

SOURCE malaria parasite P. falciparum.

ORGANISM Plasmodium falciparum

REFERENCE 1 (bases 1 to 487)

AUTHORS Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,  
 Marra,M., Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B.,  
 Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,  
 Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,  
 Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,  
 Waterston,R., Wilson,R. and Sibley,D.  
 WashU Plasmodium EST project

Unpublished (2001)

CONTACT: L. David Sibley

WashU Plasmodium EST Project

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library was constructed by Debopam Chakrabarti DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: L. David Sibley

(sibley@borcim.wustl.edu), Washington University

Seq primer: -400P from Gibco

High quality sequence stop: 384.

Location/Qualifiers

1..487

/organism="Plasmodium falciparum"

/db\_xref="taxon:5833"

/lab\_host="DH10B (Genesig, Invitrogen, Inc.)"

/note="Vector: pBluescript SK plus; Site:1: EcoRI; Site:2:

XhoI; Library was constructed by Debopam Chakrabarti.

Total RNA samples were isolated from mixed stage

saponin(0.1%) lysed P. falciparum 3D7 infected

erythrocytes by the acidic guanidinium-phenol chloroform

method. The poly A+ RNA was isolated by the polyAT-tract

mRNA isolation system (Promega, WI) using streptavidin

Magnisphere particles. Directional cDNA libraries were

constructed by oligo d(T) priming of poly(A)+ RNA (5mg)

into EcoRI and XhoI sites of 1 ZapII vector using the Zap

cDNA synthesis kit (Stratagene, CA). The average size of the

cDNA inserts in the library was between 1.0 and 1.5kb.

Clones were mass excised using the ExAssist helper phage

(Stratagene), the phagemids were precipitated with PEG

8000 and extracted with phenol/chloroform. Phagemid DNA

was electroporated into DH10B cells."

237 a 38 c 49 g 163 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 82.00 Length: 101  
 Ratio: 1.242 Gaps: 4  
 Percent Similarity: 65.347 Percent Identity: 28.713  
 alignment\_block:  
 US-09-528-682-4 x BI813876 ..  
 Align seg 1/1 to: BI813876 from: 1 to: 487

16 AsnThrAsnThrThrThrGlnPheLeuAsnMetGly.....AsnGluVa 30

175 AATACCAATATGATACAAATTTTTTAAUGTGACACAGTCTGTAATAATGA 224

30 lAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuLeuSer.....A 44

225 TATGTTACATAACAATAATAATTTTAAATTTTAAATATAATAAAG 274

44 laGluGlyGlyMetaspArgPheLeuPheGlySerAlaAsnIleAspGly 60

275 TACATAATAATATATCCAAATTTCTTTATTT...AGCAATAATATGAGTAAT 321

61 PheProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSe 77

322 TATAATGTAGTAAGTGAATAATGTGAACATAATAATAATAGTAAACAAA 371

77 rAspThrAlaSerLeuAsn.....LysGlnHisAspAlaAsp 90

372 TAATACAGCAATAATAATTTTATGATAAAAAAATAGCAACCAAT 421

90 heLysLysTyrIleLysLeuLeuLeuAsnAspGlyPhePheSerAsn 106

422 TAAATGTATTTTAAATCAAAGCTTAATAATAACTTAAATTTAAACCTAAT 471

107 Asn 107

472 AAT 474

seq\_name: gb\_est1:BE061488

seq\_documentation\_block:

LOCUS BE061488 471 bp mRNA linear EST 09-JUN-2000

DEFINITION MR0-BT0248-091299-103-e02 BT0248 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE061488

VERSION BE061488.1 GI:8406138

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 471)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR0-BT0248-091

299-103-e02&t3=1999-12-09&t4=1)

Seq primer: puc 18 forward



High quality sequence stop: 471.

## FEATURES

```

1. .471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0248"
/dev_stage="Adult"
/note="Organ: breast; Vector:
SmaI; A mini-library was made
from ORESTES PCR (U.S. Lett
,716 - Ludwig Institute for
into the pUC 18 vector. Rev
mRNA and cDNA amplification
stringency conditions."

```

Reagent Resource Center, ATCC  
<http://www.malaria.mr4.org/mr4pages/index.html>  
 Seq primer: ADF.

## FEATURES

```

FEATURES
    source
        Location/Qualifiers
            1..720
                /organism="Plasmodium yoelii yoelii"
                /strain="17XL"
                /db_xref="taxon:73239"
                /clone="pYCPM42"
                /clone_lib="PyBS"
                /dev_stage="Asexual blood stages"
                /lab_host="E. coli XL-1 Blue"
                /note="Vector: pAD-GAL4; At 20-25% parasitemia, blood was
                collected from BALB/cBYJ mice infected with Py17XL
                parasites, and leukocytes removed by passage over
                microcrystalline cellulose columns. Total RNA was
                isolated using the guanidium isothiocyanate method, and
                mRNA isolated using oligo(dT)-cellulose chromatography.
                First strand cDNA synthesis was completed using a 50-base
                primer and reverse transcriptase in the presence of
                5-methyl dCTP. After second strand synthesis, uneven
                termini were treated with Pfu DNA polymerase and EcoRI
                adaptors ligated to the blunt ends. The sample was cleaved
                with XhoI and separated on a Sephacryl S-500 column.
                Size-fractionated cDNA was precipitated and ligated to
                HybriZAP arms directionally using EcoRI-XhoI cleaved arms.
                After packaging, the phagemid vector (pAD-GAL4) was
                excised from the HybriZAP vector and plasmid DNA
    
```



```

sequence.
ACCESSION      AL275743
VERSION        AL275743.1 GI:8009945
KEYWORDS       GSS: genome survey sequence.
SOURCE         Tetraodon nigroviridis.
ORGANISM       Tetraodon nigroviridis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
               Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
               Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 1055)
AUTHORS       Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
               Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
               Weissenbach,J.
TITLE         Characterization and repeat analysis of the compact genome of the
               freshwater pufferfish Tetraodon nigroviridis
JOURNAL
REFERENCE      2 (bases 1 to 1055)
AUTHORS       Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
               Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
               Saurin,W. and Weissenbach,J.
TITLE         Human gene number estimate provided by genome wide analysis using
               Tetraodon nigroviridis DNA sequence
JOURNAL
REFERENCE      3 (bases 1 to 1055)
AUTHORS       Direct Submission
TITLE         Submitted (12-APR-2000) to the EMBL/GenBank/DBDJ databases
JOURNAL        This sequence is a single read and was generated as part of a large
COMMENT        scale clone-end sequencing project of the Tetraodon nigroviridis
               genome. For more information, please take a look at
               http://www.genoscope.cns.fr/Tetraodon.
FEATURES       Location/Qualifiers
               1..1055
               /organism="Tetraodon nigroviridis"
               /db_xref="taxon:99883"
               /clone_lib="084A20"
               /clone_lib="G"
               /note="Genoscope sequence ID : COBG084BAL0SP1-end :
               PUC-ori"
BASE COUNT     272 a 247 c 264 g 248 t 24 others
ORIGIN
alignment_scores:
Quality:       78.00      Length: 109
Ratio:         1.345      Gaps: 7
Percent Similarity: 53.211 Percent Identity: 30.275
alignment_block:
US-09-528-682-4 x CNS045RA ..
Align seg 1/1 to: CNS045RA from: 1 to: 1055
15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGly.....AsnGly 29
   ::::::::::::::: ::::::::::: :::: ::::
692 TCAAAACACCGTCCGCTCTCAGATGCCGACCTCGGCGCTTGGGGAG 741
29 uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuSerAlaGluG 46
   ::::::::::::::: ::::::::::: :::: ::::
742 CATCCGCTGGGAGGGAGG...CGCACSSTAGGCATCCCTCTCCGCTG 788
46 lYGlyMetAsp.....ArgAspLeu.....PheGlySerAla 56
   ::::::::::::::: ::::::::::: :::: ::::
789 GGGGAAGACCTGCAGCTCAGGACCTTCGTAAACTGGGGAGCGCA 838
57 AsnIleAspGlyPhe.....ProGluValArgIlePheAsnSe 69
   ::::::::::::::: ::::::::::: :::: ::::
839 AACACGCGTGTATTAAATAGCCGCGCGCTTGAAAGATCACATAG 888
69 rIeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
   ::::::::::::::: ::::::::::: :::: ::::
889 C.....CCGCAGC 896

```

```

86 isAspAlaAsp.....PheLysLysTyrIleLysLeuLeuIleAsn 99
   ::::::::::::::: ::::::::::: :::: ::::
897 ACGAGAATAGCTTCATTAGTTTAAAAAACAACACTGCTTGATCAAT 946
100 AsnAspGlyPhePheSerAsnAsnGly 108
   ::::::::::::::: ::::::::::: :::: ::::
947 AATAAATCGTCATTTBATCAAAATGGC 973
seq_name: gb_gss:CNS06FOV

seq_documentation_block:
LOCUS      CNS06FOV
DEFINITION T3 end of clone AR0AA031H06 of library AR0AA from strain CBS 732 of
            Zygocacccharomyces rouxii, genomic survey sequence.
ACCESSION  AL396677
VERSION    AL396677.1 GI:12148934
KEYWORDS   GSS.
SOURCE     Zygocacccharomyces rouxii.
ORGANISM   Zygocacccharomyces rouxii
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Zygocacccharomyces.
REFERENCE  1 (bases 1 to 982)
AUTHORS    de Montigny,J., Straub,M., Potier,S., Tekalia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Souciet,J.
TITLE      Genomic exploration of the hemiascomycetous yeasts: 8.
JOURNAL    FEBS Lett. 487 (1), 52-55 (2000)
MEDLINE    20584718
AUTHORS     Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekalia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
TITLE      Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
JOURNAL    FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE    20584711
AUTHORS     Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygocacccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES     Location/Qualifiers
              1..982
              /organism="Zygocacccharomyces rouxii"
              /strain="CBS 732"
              /db_xref="taxon:4956"
              /clone_lib="AR0AA031H06"
              /clone_lib="AR0AA"
              /note="end : T3"
              <2..>981
              /note="similar to Saccharomyces cerevisiae ORF YDL195w [
              SEC31 : component of the COPII coat of ER-golgi vesicles
              ]
              2 putative frameshift(s)"
              /evidence-not_experimental
              313 a 196 c 205 g 266 t 2 others
BASE COUNT
ORIGIN

```







```

FEATURES
  source      Location/Qualifiers
1..412
  /organism="Drosophila melanogaster"
  /db_xref="taxon:7227"
  /clone="AT10383"
  /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
  /sex="male"
  /dev_stage="0-3 day old Ore-R males"
  /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
  /note="Organ: ADULT testes; Vector: pOTB7; Site.1: EcoRI;
Site.2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      129 a      89 c      98 g      96 t
ORIGIN
alignment_scores:
  Quality: 74.00      Length: 85
  Ratio: 1.423      Gaps: 4
Percent Similarity: 61.176      Percent Identity: 27.059
alignment_block:
US-09-528-682-4 x BF496463 ..
Align seg 1/1 to: BF496463 from: 1 to: 412
22 GlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerAs 38
|||||:|||||:|||||
169 CAATCTCTTGACATTGGAAT..... 189
38 nTyrAlaLeuLeuSerAlaGluGlyGlyMetAspArgAspLeuPheGlyS 55
|||||:|||||:|||||:|||||:|||||
190 ....GCACTAACCTGTCGGGATGATGGCTGGATGATGACTCTCTTCGTG 235
55 erAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLeuPro 71
|||||:|||||:|||||:|||||:|||||
236 ATTGTGAGGATGAC.....GAGGAAGTAAAAAAGTTCAAAAGTATGGAT 279
72 AsnAsnLysAlaSerSerAsp.....ThrAlaSerLeuAsnLysG1 85
|||||:|||||:|||||:|||||:|||||
280 GATGCTGAAGCAACGACGCGGATCTCTGCCAATCTCTTAAAGCATT 329
85 nHisAspAlaSpheLysLysTyrIleLysLeuLeuIleAsnAsnAsp 101
|||||:|||||:|||||:|||||:|||||
330 CATCGACAAGGAAGTGAAGAAATCTTGGGAGGAGGCTCAATACGGAA 379
102 Gly 102
|||
380 GGC 382
seq_name: gb_est2:BF299368
seq_documentation_block:
LOCUS      BF299368      645 bp      mRNA      linear      EST 04-MAY-2001
DEFINITION      067PbF01 Pb cDNA #20, Charles Yowell and Jane Carlton Plasmodium
                berghei cDNA 3', mRNA sequence.
ACCESSION      BF299368
VERSION        BF299368.1 GI:13953117
KEYWORDS
SOURCE
  ORGANISM
    plasmodium berghei.
    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
  1 (bases 1 to 645)
REFERENCE
  Carlton,J.M.-R. and Dame,J.B.
  TITLE      The Plasmodium vivax and P. berghei gene sequence tag projects
  JOURNAL    Parasitol. Today 16 (10), 409 (2000)
  CONTACT    Contact: Dame JB
  COMMENT    Department of Pathobiology, College of Veterinary Medicine
            University of Florida
1015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611-0880, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20)forward.
Location/Qualifiers
1..645
  /organism="Plasmodium berghei"
  /strain="ANKA clone 15cyl (clone of the ANKA 8417 HP clone
)"
  /db_xref="taxon:5821"
  /clone_lib="Pb cDNA #20, Charles Yowell and Jane Carlton"
  /dev_stage="asynchronous blood stage"
  /lab_host="Swiss white mice"
  /note="Vector: pBluescript II vector DNA, excised from
Lamda ZAP II.; Site.1: EcoRI; Site.2: XhoI; Total RNA was
extracted from asynchronous blood stage forms of the
cloned ANKA isolate of P. berghei grown in laboratory
Swiss white mice. Contaminating host white cells had
previously been removed using a novel biomagnetic bead
protocol (J. Carlton et al., manuscript in preparation).
PolyA+ RNA was extracted and reverse transcribed using an
oligo dT-XhoI primer. Second strand cDNA was prepared
using RNase H and DNA polymerase I. EcoR I adaptors were
ligated to the cDNA, and it was digested with XhoI.
Fragments were size selected, and those between 1-5 kb
ligated into EcoRI /XhoI digested vector."
BASE COUNT      227 a      83 c      58 g      277 t
ORIGIN
alignment_scores:
  Quality: 74.00      Length: 103
  Ratio: 1.345      Gaps: 7
Percent Similarity: 53.398      Percent Identity: 29.126
alignment_block:
US-09-528-682-4 x BF299368/rev ..
Align seg 1/1 to reverse of: BF299368 from: 1 to: 645
14 IleAlaAsnThrAsn...ThrThrThrGlnPheLeu..... 24
|||||:|||||:|||||:|||||:|||||
472 ATAATCCACACCAATTTTGAAGCACAATTTTATGAGCAATTCGCAGA 423
25 .....AsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerA 38
|||||:|||||:|||||:|||||:|||||
422 ACTTTTAAAAATGTAAAAAATGAAAATGAT...TCGGTGATTATCCCA 376
38 snTyrAlaLeuLeuSerAlaGluGlyGlyMetAspArgAspLeuPheGly 54
|||||:|||||:|||||:|||||:|||||
375 ATTATTATATATAATAATGAAGCCGAGTGAAGGAA..... 338
55 SerAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLeuPr 71
|||||:|||||:|||||:|||||:|||||
337 .....TCTATAGACAGTTTTAAATTTTATAAC 312
71 oAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspA 88
|||||:|||||:|||||:|||||:|||||
311 CATATATAAA.....GAAGATATAATAAGATAAACAACAACAAG 271
88 laAspPheLysLysTyrIleLysLeuLeuIle.....AsnAsnAspGly 102
|||||:|||||:|||||:|||||:|||||
270 CCGAT...AAAAAATATATTATTCAATTAGTACCATCTAAAACTCAAAAT 224
103 PhePheSer 105
|||||:|||||:|||||:|||||:|||||
223 TTTTATAAT 215
seq_name: gb_est2:BF293289
seq_documentation_block:
LOCUS      BF293289
752 bp      mRNA      linear      EST 18-OCT-2001

```







```
alignment_block:
US-09-528-682-4 x BH302927

Align seg 1/1 to: BH302927 from: 1 to: 816

3 pheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnTh 19
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 TTCACACACAGAGATACAGAGATGGAAGAGAGATCTCAGGAACAGAAGA 285

19 rThrThrGlnPheLeuAsnMetGly..... 27
::: ::::::::::::::::::::

286 TTCCATAGAAATCATTTGACTCACTGCTCAAGATATATGTAAGCGGAAAA 335

28 .....:::AsnGluValAlaLeuAspGlyArgTyrSer 37
::: |||||::: ::: |||||

336 AGCTACTGCTCAAAACATACAGGAATCCAGGACTCAATGAGAGATCA 385

38 AsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG1 54
||| ::|||::: ||| ::||| ::||| ||| ::||| |||
386 AACCTAAGGATATAGGTATAGTAGAGAGTGAAGACTCCCGCTCAAAGG 435

54 ySerAlaAsnIle.....AspGlyPhePro..... 62
| ::||| ||| ::::: |||||

436 ACCAGTAAATATCTTCAACAAATCATGAGAAACCTCCCTAACCTAA 485

63 .....GluValArgGluPheAsnSerLeuProAsnAsn 73
::: ::|||::: ||| ::|||:::

486 AAAAGAGATACCCATAGACATACAGAAGCCTACAGAACTCCAATAGA 535

74 LysAlaSerSerAspThrAlaSer.....LeuAsnLysGlnHisAs 87
::: ::|||::: ||| ::|||::: ||| ::|||:::

536 AACACCAAGAAAGAAACACCCCGCTCACATAATTTGTTGAAAAAACACA 585

87 pAlaAspPheLysLysTyrIleLys 95
::: ||| |||||

586 AACGCACAAAAATAAGATATTAA 610

seq_name: gb_gss:AZ930886

seq_documentation_block:
LOCUS AZ930886 501 bp DNA linear GSS 01-APR-2001
DEFINITION 474.dhz59f11.sl Saccharomyces unisporus NRRL Y-1556 Saccharomyces
unisporus genomic clone 474.dhz59f11.sl, DNA sequence.
ACCESSION AZ930886
VERSION AZ930886.1 GI:13501796
KEYWORDS GSS.
SOURCE Saccharomyces unisporus.
ORGANISM Saccharomyces unisporus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE
1 (bases 1 to 501)
Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
W.R., Waterston,R.H. and Johnston,M.
Contact: Johnston M
TITLE Surveying Saccharomyces genomes to identify functional elements by
comparative DNA sequence analysis
JOURNAL Unpublished (2001)
COMMENT Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mj@genetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..501
/organism="Saccharomyces unisporus"
/strain="NRRL Y-1556 (CBS 398)"
/db_xref="taxon:27294"
/clone="474.dhz59f11.sl"
/clone_lib="Saccharomyces unisporus NRRL Y-1556"
/note="Random genomic sequence"
BASE COUNT 151 a 86 c 72 g 192 t
```

```
ORIGIN

alignment_scores:
Quality: 73.50 Length: 116
Ratio: 1.205 Gaps: 6
Percent Similarity: 52.586 Percent Identity: 28.448

alignment_block:
US-09-528-682-4 x AZ930886/rev ..

Align seg 1/1 to reverse of: AZ930886 from: 1 to: 501

16 AsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLe 32
||| ||||| ||||| ||||| ||||| ||||| |||||
501 AACAAACACACTATTACAAAT...CTAAATTTAGTTCATGTTAACT.. 457

32 uAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetA 49
||| ::|||::: ||| ::||| ::||| ||| ::||| |||
456 .GATAAAGATATACACATTAT.....G 435

49 sPArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArg 65
||||| |||||::: ||| ::|||::: ||| ::|||:::

434 ATAGACCGCTTATATCGAGTCCCAATCTGCAATGATAGTCAAAATTATA 385

66 GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLe 82
||| ::|||::: ||| ::|||::: ||| ::|||:::

384 GAATAAATGCTTACCAGATGATGATAATGAAGCAGAGAGAGGACCCA 335

82 uAsnLysGlnHisAspAlaAsp...Phe..... 90
::: ::|||::: ||| ::|||::: ||| ::|||:::

334 CGATAGGCATCACTCTCGACATAGTTTCACCTGAATTAAGAGAGACCAAA 285

91 .....LysLysTyrIleLysLeuL 97
::: ::|||::: ||| ::|||::: ||| ::|||:::

284 GCTCTGGAAGGGCCTTTTAGAATCATCACCGCAATGATGATCAAGTTAT 235

97 euille.....AsnAsnAspGlyPhePheSerAsnAsnGlyGlyLys 110
||| ::|||::: ||| ::|||::: ||| ::|||:::

234 AAATATGTGGTGAATATGATGGTATACTTTCGAATAATAAGAGAAAA 188

seq_name: gb_gss:AZ523935

seq_documentation_block:
LOCUS AZ523935 664 bp DNA linear GSS 07-MAY-2001
DEFINITION 225PbC02 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.
ACCESSION AZ523935
VERSION AZ523935.1 GI:13963369
KEYWORDS GSS.
SOURCE Plasmodium berghei.
ORGANISM Plasmodium berghei.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 664)
Carlton,J.M.-R. and Dame,J.B.
AUTHORS The Plasmodium vivax and P. berghei gene sequence tag projects
TITLE Parasitol. Today 16 (10), 409 (2000)
JOURNAL Contact: Dame JB
COMMENT Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.
FEATURES
source
1..664
/organism="Plasmodium berghei"
/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
/db_xref="taxon:5821"
/clone_lib="Pb MBN #21"
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
```







```

seq_name: gb_est1:AW257463
seq_documentation_block:
LOCUS      AW257463                655 bp    mRNA    linear    EST 20-DEC-1999
DEFINITION EST305600 KV2 Medicago truncatula cDNA clone KV2-9L18, mRNA
sequence.
ACCESSION  AW257463
VERSION    AW257463.1 GI:6605720
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 655)
AUTHORS   VandenBosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,
            Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
            Fraser,C.M.
TITLE     ESTs from roots of Medicago truncatula after Rhizobium inoculation
JOURNAL   Unpublished (1999)
COMMENT   Contact: VandenBosch K
            Department of Biology
            Texas A&M University
            College Station, TX 77843-3258, USA
            Tel: 409 845 7707
            Fax: 409 845 2891
            Email: kate@mail.bio.tamu.edu
            Texas A&M EST name: r1l6168e
            TIGR sequence name: MTAAT69TK
            More information is available at: . (and for clone ordering info)
            http://chryslie.tamu.edu/medicago
            Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
FEATURES   Location/Qualifiers
            source
            1..655
            /organism="Medicago truncatula"
            /cultivar="genotype Al7"
            /db_xref="taxon:3880"
            /clone_lib="KV2-9L18"
            /clone_lib="KV2"
            /tissue_type="Seedling roots"
            /dev_stage="2 days post-inoculation with Sinorhizobium
            meliloti"
            /lab_host="E. coli strain SOLR"
            /note="Vector: pBluescript SK -; Site.1: EcoRI; Site.2:
            XhoI; cDNA was prepared from polyA+ enriched RNA. The
            cDNA was directionally ligated into the Unizap XR vector
            from Stratagene and packaged using Gigapack III Gold
            packaging extracts. Plasmids containing cDNA inserts
            were excised from the recombinant lambda-zap phage using
            Ex-assist helper phage and propagated in SOLR cells."
BASE COUNT      220 a 138 c 109 g 188 t
ORIGIN
18 AsnThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspG1 34
|||||
328 AATGTAATAAATCAAGTACATCAGTGTGGAAATGAATTAACCACTAATGA 377
|||||
34 yAGTgYrSerAsnTyAlaLeuIleSerAlaGluGlyGlyMetAspArgA 51
|||||
378 CAATGAGGCCCAATACATCTCTCCCTGCCATGCAAAAC.....ATTCAA 421

```

```

51 spLeuPheGlySerAlaAsnIleAspGly..... 60
52 ATGCAATTTTCATCTGCAAAATTTTCAAGGCCAAATCAAGGTCTCAACAGCA 471
61 .....PheProGluValIArgGluPheAs 68
472 ATAGACATGACTTTGATTGGTAAGTCAATTTCCA..... 504
68 nSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysG 85
505 .....CCCAATGATGGTGTGTTTCAGTGACCAAGCA..... 534
85 InHisAspAlaAspPheLysLysTyIleLysLeuLeuIleAsnAsnAsp 101
535 .....AAACCATATATACACCAATAATAAC..... 561
102 GlyPhePheSerAsnAsnGly 108
562 ...TTCTTAACAACAATGGA 579
seq_name: gb_est2:BG647231
seq_documentation_block:
LOCUS      BG647231                736 bp    mRNA    linear    EST 24-APR-2001
DEFINITION EST508850 HOGA Medicago truncatula cDNA clone PHOGA-16E21 5' end,
mRNA sequence.
ACCESSION  BG647231
VERSION    BG647231.1 GI:13782343
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 736)
AUTHORS   Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
            Utterback,T., Cho,J. and Fraser,C.M.
TITLE     ESTs from roots of Medicago truncatula treated with
            oligogalacturonides of DP 6-20
JOURNAL   Unpublished (2001)
COMMENT   Contact: Michael G. Hahn
            Complex Carbohydrate Research Center
            University of Georgia
            220 Riverbend Road, Athens, GA 30602-4712, USA
            Tel: 706-542-4457
            Fax: 706-542-4412
            Email: hahn@ccrc.uga.edu
            G390837e TIGR sequence name: MTMBW35TK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
FEATURES   Location/Qualifiers
            source
            1..736
            /organism="Medicago truncatula"
            /cultivar="Al7"
            /db_xref="taxon:3880"
            /clone_lib="PHOGA-16E21"
            /clone_lib="HOGA"
            /tissue_type="3 day old seedling roots"
            /dev_stage="24 hours after treatment in the dark at 26 C
            with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
            presence of 100 ug/ml Gentamicin"
            /lab_host="XL0LR"
            /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
            XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
            was directionally ligated into the Unizap XR vector from
            Stratagene and packaged using Gigapack III Gold packaging
            extracts. Plasmids containing cDNA inserts were excised
            from the recombinant lambda-zap phage using Ex-assist
            helper phage and propagated in SOLR cells."
BASE COUNT      245 a 157 c 123 g 211 t
ORIGIN

```



```
alignment_scores:
  Quality: 72.50      Length: 107
  Ratio: 1.394      Gaps: 5
  Percent Similarity: 48.598      Percent Identity: 24.299

alignment_block:
  US-09-528-682-4 x BG647231 ..
  Align seq 1/1 to: BG647231 from: 1 to: 736
  18 AsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspG1 34
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  328 AATGTAATAATCAAGTACATCATCTTCCTGCGATGCAAAAC.....ATTCAAA 421
  34 YAGTyrSerAsnTyrAlaLeuLeuSerAlaGluGlyGlyMetAspArg 51
  ::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  378 CAATGAGGCCCAATACATCTTCCTGCGATGCAAAAC.....ATTCAAA 421
  51 spLeuPheGlySerAlaAsnIleAspGly..... 60
  422 ATGCAATTCATCTGCAATTTTCAAGGCCCAATCAAGGTCTCAACAGCA 471
  61 .....PheProGluValArgGluPheAs 68
  472 ATAGACATGACTTTGATGTTAGTCATTTCCA..... 504
  68 nSerLeuProAsnLysAlaSerSerAspThrAlaSerLeuAsnLysG 85
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  505 .....CCCAATGATGGTGTTCCTGATGACCAAGCA..... 534
  85 lnHisAlaAspPhePheLysTyrIleLysLeuLeuIleAsnAsnAsp 101
  ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
  535 .....AAACCATATATACACCAATAATAAAC..... 561
  102 GlyPhePheSerAsnAsnGly 108
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  562 ...TTCTAAACAACAATGGA 579

seq_name: gb_gss:BH486660

seq_documentation_block:
  LOCUS BH486660 773 bp DNA linear GSS 13-DEC-2001
  DEFINITION BOHKF67R BOHK Brassica oleracea genomic clone BOHKF67, DNA
  sequence.
  ACCESSION BH486660.1 GI:17694764
  VERSION BH486660
  KEYWORDS GSS.
  SOURCE Brassica oleracea.
  ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
  REFERENCE 1 (bases 1 to 773)
  AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL Unpublished (2001)
  COMMENT Other_GSSs: BOHKF67TF
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TR
  Class: sheared ends.
  FEATURES
    source
      1..773
        /organism="Brassica oleracea"
        /strain="T01000DH3"
        /db_xref="taxon:3712"

alignment_scores:
  Quality: 72.50      Length: 107
  Ratio: 1.394      Gaps: 5
  Percent Similarity: 48.598      Percent Identity: 24.299

alignment_block:
  US-09-528-682-4 x BG647231 ..
  Align seq 1/1 to: BG647231 from: 1 to: 736
  18 AsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspG1 34
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  328 AATGTAATAATCAAGTACATCATCTTCCTGCGATGCAAAAC.....ATTCAAA 421
  34 YAGTyrSerAsnTyrAlaLeuLeuSerAlaGluGlyGlyMetAspArg 51
  ::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  378 CAATGAGGCCCAATACATCTTCCTGCGATGCAAAAC.....ATTCAAA 421
  51 spLeuPheGlySerAlaAsnIleAspGly..... 60
  422 ATGCAATTCATCTGCAATTTTCAAGGCCCAATCAAGGTCTCAACAGCA 471
  61 .....PheProGluValArgGluPheAs 68
  472 ATAGACATGACTTTGATGTTAGTCATTTCCA..... 504
  68 nSerLeuProAsnLysAlaSerSerAspThrAlaSerLeuAsnLysG 85
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  505 .....CCCAATGATGGTGTTCCTGATGACCAAGCA..... 534
  85 lnHisAlaAspPhePheLysTyrIleLysLeuLeuIleAsnAsnAsp 101
  ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
  535 .....AAACCATATATACACCAATAATAAAC..... 561
  102 GlyPhePheSerAsnAsnGly 108
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  562 ...TTCTAAACAACAATGGA 579

seq_name: gb_gss:BH486660

seq_documentation_block:
  LOCUS BH486660 773 bp DNA linear GSS 13-DEC-2001
  DEFINITION BOHKF67R BOHK Brassica oleracea genomic clone BOHKF67, DNA
  sequence.
  ACCESSION BH486660.1 GI:17694764
  VERSION BH486660
  KEYWORDS GSS.
  SOURCE Brassica oleracea.
  ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
  REFERENCE 1 (bases 1 to 773)
  AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL Unpublished (2001)
  COMMENT Other_GSSs: BOHKF67TF
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TR
  Class: sheared ends.
  FEATURES
    source
      1..773
        /organism="Brassica oleracea"
        /strain="T01000DH3"
        /db_xref="taxon:3712"
```

```
/clone="BOHKF67"
/clone_lib="BOHK"
/note="Vector: phos1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into phos1 using BstXI linkers"
BASE COUNT 222 a 147 c 207 g 197 t
ORIGIN

alignment_scores:
  Quality: 72.50      Length: 59
  Ratio: 2.071      Gaps: 2
  Percent Similarity: 59.322      Percent Identity: 35.593

alignment_block:
  US-09-528-682-4 x BH486660 ..
  Align seq 1/1 to: BH486660 from: 1 to: 773
  9 AspAlaTyrGluProIle.....AlaAsnThrAsnThrThrThrGlnPh 23
  |||::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  383 GATCTGGTACACCTGTAACACAAGCTAATACAATGAGGAGCCCCAAGC 432
  23 eLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrA 40
  ::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  433 CATAAGCTTGGAAATATTTCCTGTTAGATGGATCTTGGACGCTCTCTG 482
  40 laLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAla 56
  || : : : : : : : : : : : : : : : : : : : : : : : : : : :
  483 CTACTTTAGTGGATGCGGATGGCGTGGATGGATGGATGGATGGGAATGCA 532
  57 AsnIleAspGly.....PhePro 62
  ::: : : : : : : : : : : : : : : : : : : : : : : : : : :
  533 CAGCTTATGGGACACAGGAATTTCCCT 559

seq_name: gb_gss:BH552001

seq_documentation_block:
  LOCUS BH552001 801 bp DNA linear GSS 14-DEC-2001
  DEFINITION BOHMW73TF BOHM Brassica oleracea genomic clone BOHMW73, DNA
  sequence.
  ACCESSION BH552001
  VERSION BH552001.1 GI:17803781
  KEYWORDS GSS.
  SOURCE Brassica oleracea.
  ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
  REFERENCE 1 (bases 1 to 801)
  AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL Unpublished (2001)
  COMMENT Other_GSSs: BOHMW73TR
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TF
  Class: sheared ends.
  FEATURES
    source
      1..801
        /organism="Brassica oleracea"
        /strain="T01000DH3"
        /db_xref="taxon:3712"
        /clone="BOHMW73"
        /clone_lib="BOHM"
        /note="Vector: phos1; Site_1: BstXI; 2-3 kb sheared
        genomic DNA inserted into phos1 using BstXI linkers"
BASE COUNT 202 a 212 c 158 g 229 t
ORIGIN
```







Quality: 72.00 Length: 121  
Ratio: 1.309 Gaps: 4  
Percent Similarity: 45.455 Percent Identity: 23.967

## alignment\_block:

US-09-528-682-4 x AV916505 ..

Align seg 1/1 to: AV916505 from: 1 to: 620

```

13 ProIleAlaAsnThrAsnThrThr.....ThrGlnPheLeuAsnMetG1 27
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
45 CTGTAGTCAACCTCTCTGCTGCGTTTGGAACTAGTCTAGCTGTGG 94
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
27 YAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerA 44
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
95 TGTGTATGTCTGCTTGATACATCCACCAAGAATTTCAACCAATACAATG 144
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
44 laGlu...GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAsp 59
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
145 CTGCGTTTAGCTACACTAACCAAGATCTTATCGCATCCCTGGAACCTG... 191
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 GlyPheProgluValArgGluPheAsnSerLeuProAsnAsnLysAlaSe 76
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
192 .....AACCAAGGGAGA 205
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
76 rSerAspThrAlaSer..... 81
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
206 CAGCCTCAGCGCATCTTATACCACTTTGTGGAGAGTCAGCACAGCTG 255
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
81 ..... 81
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
256 TTGGGGCGGAGCTGACCCACAGCTTCTCGAGCAACGAGACAGCCTCACC 305
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
82 LeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLeuLeuII 98
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
306 TTGGGCACTCAGCATACCTTAGACCCGCTCACCCTCGTGAAGCTCGCAT 355
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 eAsnAsnAspGly 102
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
356 CAACAACCTCTGGC 368
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

seq\_name: gb\_estl:AV916629

## seq\_documentation\_block:

LOCUS AV916629 633 bp mRNA linear EST 18-JAN-2002  
DEFINITION AV916629 K. Sato unpublished cDNA library, cv. Haruna NiJo  
germination shoots Hordeum vulgare subsp. vulgare cDNA clone  
bags19b07 5', mRNA sequence.

ACCESSION AV916629

VERSION AV916629

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 633)

AUTHORS Sato, K., Saisho, D. and Takeda, K.

TITLE Barley EST sequencing project in NIG and Okayama Univ

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

FEATURES

source

1..633

/organism="Hordeum vulgare subsp. vulgare"

/cultivar="Haruna NiJo"

/db\_xref="taxon:112509"

/clone="bags19b07"

/clone\_lib="K. Sato unpublished cDNA library, cv. Haruna  
NiJo germination shoots"  
/tissue\_type="shoots"  
/dev\_stage="germination"

BASE COUNT 157 a 180 c 150 g 145 t 1 others  
ORIGIN

## alignment\_scores:

Quality: 72.00 Length: 121  
Ratio: 1.309 Gaps: 4  
Percent Similarity: 45.455 Percent Identity: 23.967

## alignment\_block:

US-09-528-682-4 x AV916629 ..

Align seg 1/1 to: AV916629 from: 1 to: 633

```

13 ProIleAlaAsnThrAsnThrThr.....ThrGlnPheLeuAsnMetG1 27
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
147 CTGTAGTCAACCTCTCTGCTGCGTTTGGAACTAGTCTAGCTGTGG 196
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
27 YAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerA 44
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
197 TGTGTATGTCTGCTTGATACATCCACCAAGAATTTCAACCAATACAATG 246
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
44 laGlu...GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAsp 59
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
247 CTGCGTTTAGCTACACTAACCAAGATCTTATCGCATCCCTGGAACCTG... 293
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 GlyPheProgluValArgGluPheAsnSerLeuProAsnAsnLysAlaSe 76
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
294 .....AACCAAGGGAGA 307
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
76 rSerAspThrAlaSer..... 81
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
308 CAGCCTCAGCGCATCTTATACCACTTTGTGGAGAGTCAGCACAGCTG 357
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
81 ..... 81
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
358 TTGGGGCGGAGCTGACCCACAGCTTCTCGAGCAACGAGACAGCCTCACC 407
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
82 LeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLeuLeuII 98
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
408 TTGGGCACTCAGCATACCTTAGACCCGCTCACCCTCGTGAAGCTCGCAT 457
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 eAsnAsnAspGly 102
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
458 CAACAACCTCTGGC 470
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

seq\_name: gb\_estl:AV917534

## seq\_documentation\_block:

LOCUS AV917534 715 bp mRNA linear EST 18-JAN-2002  
DEFINITION AV917534 K. Sato unpublished cDNA library, cv. Haruna NiJo  
germination shoots Hordeum vulgare subsp. vulgare cDNA clone  
bags10f08 3', mRNA sequence.

ACCESSION AV917534

VERSION AV917534

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 715)

AUTHORS Sato, K., Saisho, D. and Takeda, K.

TITLE Barley EST sequencing project in NIG and Okayama Univ

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan











[illegible]







```

seq_documentation_block:
LOCUS      AZ006675                624 bp    DNA        linear    GSS 25-FEB-2000
DEFINITION RPCI-23-31511.TJ RPCI-23 Mus musculus genomic clone RPCI-23-31511,
            DNA sequence.
ACCESSION  AZ006675
VERSION    AZ006675.1 GI:7082059
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 624)
AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE      Mouse BAC End Sequences from Library RPCI-23
JOURNAL    Unpublished (1999)
COMMENT    Other_GSSs: RPCI-23-31511.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact pletier de Jong
            (pletier@jeong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 315 row: I column: 1
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..624
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-31511"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site:1;
                     EcoRI; Site:2; EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and partially digested
                     with a combination of EcoRI and EcoRI Methylase. Size
                     selected DNA was cloned into the pBACE3.6 vector at the
                     EcoRI sites. The ligation products were transformed into
                     DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT  299 a 113 c 108 g 104 t
ORIGIN
alignment_scores:
Quality: 71.50 Length: 107
Ratio: 1.172 Gaps: 6
Percent Similarity: 57.009 Percent Identity: 31.776
alignment_block:
US-09-528-682-4 x AZ006675
Align seg 1/1 to: AZ006675 from: 1 to: 624
6 ArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnThrThr.. 21
143 AGAATCTCAGAGCGAGAAGATACCATGAGAAACATTGACACAAATGCAA 192
22 .....GlnPheLeuAsnMetGly...AsnGluVala 31
193 AGAAATGTAAATGCAAAAGATTCTTCCACCAAAACATCCAGGAATCC 242

```

```

31 laLeuAspGlyArgTyrSerAsnTyrAlaLeuLeuSerAlaGluGlyGly 47
    ||| ||| ::||| ||| ||| |||
243 AGGACAAATGAGAAGACCAACCTAAGGATAATAAGTATAGAAGGAGT 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 MetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVa 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
293 AAAGATTCCCAACTTAAGGGACCCAGTAATATC.....TTCAACAATA 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 laIrgGlu.....PheAsnSerLeuProAsnAsnLysAlaSerSerAsp 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 TATAGAGAAAACCTTCCTACCTAAATAAAAGAGATGCCCATTAACATA 386
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
79 hrAlaSerLeu...AsnLysGlnHisAspAlaAspPheLysLysTyrIle 94
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 CAAGAAGCCTACAGAACTCCAAATAGATTGTACCAGTAAGAATAATTC... 433
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
95 LysLeuLeuIleAsnAsnAsp 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 ...CTCCTGTGCACATAATAAT 451

seq_name: gb_est2:BE366639

seq_documentation_block:
LOCUS      BE366639                666 bp    mRNA        linear    EST 20-JUL-2000
DEFINITION P11_35_D02.g1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
            mRNA sequence.
ACCESSION  BE366639
VERSION    BE366639.1 GI:9308196
KEYWORDS   EST.
SOURCE     sorghum.
ORGANISM   Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE  1 (bases 1 to 666)
AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
            ,L.H.
TITLE      An EST database from Sorghum: pathogen-induced plants
JOURNAL    Unpublished (2000)
COMMENT    Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 12
            High quality sequence stop: 666
            POLYA-No. Location/Qualifiers
            1..666
            /organism="Sorghum bicolor"
            /db_xref="taxon:4558"
            /clone_lib="Pathogen induced 1 (P11)"
            /note="Organ: Anthracnose-infected leaves from
            two-week-old sorghum plants 48 hr after inoculation;
            Vector: pBluescript II from Lambda Zap II; Site:1: XhoI;
            Site:2: EcoRI; Two-week-old sorghum plants (B7X 623
            cultivar) were infected with pathogen (isolate FRM421 of
            Colletotrichum graminicola, which is a sorghum isolate).
            RNA was prepared from infected leaves harvested from 45
            seedlings 48 hours after inoculation. Note: young
            seedlings (2 weeks old) exhibit juvenile resistant
            reaction, which is an incompatible interaction. As they
            grow older (4 weeks or older), plants resume susceptibility
            to anthracnose disease. The library was made from poly-A
            RNA in the cloning vector lambda Zap II. Clones to be
            sequenced were prepared by mass excision. WARNING: While
            most or all ESTs are expected to derive from the host

```







FEATURES	source
1. .550	
Location/Qualifiers	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone_lib="CT0220"	
/dev_stage="Adult"	
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
101 a	134 c 179 g 136 t
BASE COUNT	
ORIGIN	

```

Align seg 1/1 to reverse of: AW377167 from: 1 to: 550

13  ProlLeuAlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnG1 29
    |||  :::::|||||:::  :::::  :::::  |||
309  CCCCGAGATTCAACAAATTCAGATGCACACATCTACACAAATGATTAATCC 260

29  uValAlaLeuAsnGlyArgTyrSerAsnTyrAlaLeu..... 41
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
259  AGTCCCGCTGGAGGCGCGGTCACCACTTCCCGATGGGTCCCGGCTCGG 210

42  .....IleSerAlaGluGlyGlyMetAsp.....ArgAsp 51
    :::::  :::::  |||  |||  |||  |||  :::
209  ACGTCCGATGGCGGCGCATGGTGGCATGGAGCCACCCACCATGAATGCA 160

52  LeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheAs 68
    :::::|||||:::|||||:::|||||:::|||||:::
159  TCATTAGGGTCAGGCGACATAGACGAGCTTCCAAA.....AA 122

68  nSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysG 85

```

[illegible]















48 eAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64  
 ::::: ::::: |||||  
 788 TTGAAGGCGTTTGACAGCGCTGTACATTCGCCGAGATCCGACTTCA 837  
 65 ArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSe 81  
 ::::: |||||  
 838 CAAGGTTTAAAT.....AGTAAAGATCAAGACTC 866  
 81 rLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysLeuLeuI 98  
 | ||||| ::::: ::::: |  
 867 CGAWACAGGTTCAATCCACAATTTCAAT.....TACA 901  
 98 leAsnAsnAsp 101  
 ||||| ::::: |  
 902 TCAATAGTGTAT 912

seq\_name: gb\_est1:AI907481

seq\_documentation\_block:

LOCUS AI907481 502 bp mRNA linear EST 30-MAR-2000

DEFINITION RC-BT144-190399-038 BT144 Homo sapiens cDNA, mRNA sequence.

ACCESSION AI907481

VERSION AI907481.1 GI:6497911

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/seq/gethtml.pl?tl=RC&t2=RC-BT144-038.html

&t3=190399&t4=1)

Seq primer: puc 18 forward.

Location/Qualifiers

FEATURES

source

1..502

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="BT144"

/sex="female"

/dev\_stage="Adult"

/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 123 a 107 c 118 g 153 t 1 others

ORIGIN

alignment\_scores:

Quality: 70.50

Length: 102

Ratio: 1.469

Percent Similarity: 47.059 Percent Identity: 24.510

alignment\_block:

US-09-528-682-4 x AI907481/rev ..

Align seg 1/1 to reverse of: AI907481 from: 1 to: 502

8 GlnAspAlaTyrGluProIleAlaAsnThrAsnThrThrGlnPheLe 24  
 ::::: |||||  
 466 GAAGATACATAT.....ACAAGCACCCCAATTTAA 432  
 24 uAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaL 41  
 ::::: |||||  
 431 CATATTGGCAATAACTCCCTCCCATTTCCCTCCCAATTTTCAA 382  
 41 euIleSerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsn 57  
 ::::: |||||  
 381 TAGTAGTGTTTTAAATAAATTAAGACATGTCACACAGGGGAAGATGGC 332  
 58 IleAspGlyPheProGluValArgGluPheAsn..... 68  
 ||| ::::: |||||  
 331 ATCTTCAATTTCTCAAAATTTACTGAGTCAGCCCTGCCCAAGGTTGTG 282  
 69 .....SerLeuProAsnA 73  
 |||||  
 281 GGAAGAAGGGGATGAGAGCCAGCAGGCAAGCCCTTCACTGCCTCCAC 232  
 73 snLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAsp 89  
 ::::: |||||  
 231 ATCAATGCGGAGAAACCTGCTGATGAACAAGAACACACACTAAGGGAT 182  
 90 PheLys 91  
 |||||  
 181 TTTAGG 176

seq\_name: gb\_est2:BE581518

seq\_documentation\_block:

LOCUS BE581518 505 bp mRNA linear EST 09-MAY-2001

DEFINITION kq47a07.yl TBN95TM-SSR Strongyloides stercoralis cDNA 5', mRNA

sequence.

ACCESSION BE581518

VERSION BE581518.1 GI:9832460

KEYWORDS EST.

SOURCE Strongyloides stercoralis.

ORGANISM Strongyloides stercoralis

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Panagrolaimoidea; Strongyloidea; Strongyloidea; Strongyloidea;

REFERENCE 1 (bases 1 to 505)

AUTHORS  
 McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,  
 Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,  
 Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R.,  
 Ronko, I., Kennedy, S., Maguire, J., Beck, C., Underwood, K., Steptoe,  
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,  
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and  
 Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999

JOURNAL

COMMENT

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

The library was constructed by Dr. Thomas Nutman and colleagues of  
 NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington  
 University Genome Sequencing Center St. Louis.

High quality sequence stop: 401.

FEATURES

source

1..505

/organism="Strongyloides stercoralis"







```

seq_name: gb_est2:BJ164638

seq_documentation_block:
LOCUS       BJ164638               680 bp      mRNA      linear      EST 24-JAN-2002
DEFINITION   BJ164638 full length cDNA library, chloronemata and young gametophores Physcomitrella patens subsp. patens cDNA clone ph2if20 3', mRNA sequence.
ACCESSION   BJ164638
VERSION     BJ164638.1 GI:18332625
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens.
ORGANISM    Physcomitrella patens subsp. patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1 (bases 1 to 680)
            Fujita,T., Shin-i,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.
TITLE       Comparison of the moss Physcomitrella patens genome with flowering plants genome
JOURNAL     Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            A backbone of the vector is basically from pBluescript(KS), that was in vivo excised from a modified lps phage vector (Mo bi Tec, Germany). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated basically according to the method described in The Plant J 15, 707-720 (1998) Seki M. et al. Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light.

FEATURES             source
     source           1..680
                     /organism="Physcomitrella patens subsp. patens"
                     /db_xref="taxon:145481"
                     /clone="pph2if20"
                     /clone_lib="full length cDNA library, chloronemata and young gametophores"
                     /tissue_type="mixture of chloronemata and young gametophores with 2 to 5 leaves"
BASE COUNT        185 a 164 c 156 g 175 t
ORIGIN

alignment_scores:
    Quality:       70.50      Length:       99
    Ratio:         1.259      Gaps:       3
    Percent Similarity: 56.566      Percent Identity: 26.263

alignment_block:
US-09-528-682-4 x BJ164638/rev ..

Align seg 1/1 to reverse of: BJ164638 from: 1 to: 680

      2  TyrPheThrValArgIleGlnAspAlaTyrGluProlleAlaAsnThrAs 18
      |||:||||| : : : : |||
      422  TACATCACAGAAATGGTATGGATGACTCGGACGACCTCGGCAACACCAA 373
      |||:||||| : : : : |||
      18  nThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyA 35
      : : : |||:||||| : : : : |||
      372  GCTTCGGACAGATTTCTAACGATTCACACGCTATCGCTTACTATGAA 323
      : : : |||:||||| : : : : |||
      35  rgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAsp...Arg 50
      : ||| : : : ||| ||| : : : : |||
      322  ACTATCTGTCACTCTTCTTGAATCCATCCATTCGAATGTCGAATGTGGCT 273

```



```

69 rLeuPro 71
   ::::|
211 TGTTCCTCC 217

seq_name: gb_gss:BH350007

seq_documentation_block:
LOCUS      BH350007              749 bp    DNA        linear    GSS 03-DEC-2001
DEFINITION CH230-198K12.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone.
            CH230-198K12, DNA sequence.
ACCESSION  BH350007
VERSION    BH350007.1  GI:17280741
KEYWORDS   GSS.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 749)
AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL   Unpublished (1999)
COMMENT   Other_GSSs: CH230-198K12.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@email.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/orering\_information.htm). BAC end
            page: http://www.tigr.org/tldb/bac\_ends/rat/bac\_end\_intro.html
            Plate: 198 row: K column: 12
            Seq primer: SP6
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..749
                     /organism="Rattus norvegicus"
                     /strain="BN/SMNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-198K12"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="vector: pTABBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SMNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT          360 a 137 c 139 g 113 t
ORIGIN

alignment_scores:
  Quality: 70.50          Length: 126
  Ratio: 1.102           Gaps: 5
  Percent Similarity: 50.794  Percent Identity: 24.603

alignment_block:
us-09-528-682-4 x BH350007  ..
Align seg 1/1 to: BH350007 from: 1 to: 749

3 PheThrValArgIleGlnAspAlaTy rGluProIleAlaAsnThrAsnTh 19
325 TTCCACCAAGACAGATACAGAGATGGAAGAGAGAGATCTCAGGACGAGA 374
19 rThrThrGlnPheLeuAsnMetGly..... 27

```

```

   ::::|
375 TTCCATAGAAATCATTGACTCAACTGTCAAAGATAATTTAAAGCAGAAAA 424
28 .....AsnGluValAlaLeuAspGlyArgTy rSer 37
   ::::|
425 AGCTACTGCTCCAAAACATACAGGAAATCCAGGACTCAATGAGAAGATCA 474
38 AsnTy rAlaLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheG1 54
   ::::|
475 ACCTTAAGATATAGTATAGAGAGAGAGTGAGAGACTCCGAGCTCAAGG 524
54 ySerAlaAsnIle.....AspGlyPheProGlu.... 63
   ::::|
525 ACCAGTAAATATCTTCAACAAAATCATAGAAGAAACTTCCTCAAGCTAA 574
64 .....ValArgGluPheAsnSerLeuProAsnAsn 73
575 AAAAGAGTTACCCATVAGCCATACAAAGAGCCCTACAGAACTCCAAATAGA 624
74 LysAlaSerSerAspThrAlaSer.....LeuAsnLysGlnHisAs 87
   ::::|
625 TTGGACCAAAAGAAACACCTCCGTCACATATTTGTCAAAACACCCAAA 674
87 pAlaAspPheLysLys....Ty rIleLys 95
   ::::|
675 CGCACAAATAAGAGAAAGATATTTAAAG 702

seq_name: gb_est2:BG603389

seq_documentation_block:
LOCUS      BG603389              360 bp    mRNA       linear    EST 14-AUG-2001
DEFINITION EST502479 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii CDNA
            clone PYCDR11, mRNA sequence.
ACCESSION  BG603389
VERSION    BG603389.1  GI:15153403
KEYWORDS   EST.
SOURCE     Plasmodium yoelii.
ORGANISM   Plasmodium yoelii
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE  1 (bases 1 to 360)
AUTHORS   Kappe,S.H.I., Gardner,M.J., Brown,S.M., Ross,J., Matuschewski,K.,
            Ribeiro,J.M., Adams,J.H., Quackenbush,J., Cho,J., Carucci,D.J.,
            Hoffman,S.L. and Nussenzweig,V.
            Exploring the transcriptome of the malaria sporozoite stage
            Proc. Natl. Acad. Sci. U. S. A. 98 (17), 9895-9900 (2001)
            Contact: Malcolm J. Gardner
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org
            Request for clones, please contact: Stefan Kappe,
            kappes01@popmail.med.nyu.edu Michael Heidelberger Division,
            Department of Pathology New York University School of Medicine.

FEATURES             Location/Qualifiers
     source           1..360
                     /organism="Plasmodium yoelii"
                     /strain="17XNL"
                     /db_xref="taxon:5861"
                     /clone="PYCDR11"
                     /clone_lib="Plasmodium yoelii sporozoite cDNA"
                     /dev_stage="sporozoites from salivary gland"
                     /lab_host="E. coli TOP10"
                     /note="Vector: pCR4; TA cloning; Plasmodium yoelii
                     sporozoite cDNA library from salivary gland sporozoites 14
                     days post-infection"
BASE COUNT          75 a 71 c 63 g 151 t
ORIGIN

alignment_scores:

```







The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end  
plate: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)  
Plate: 113 row: M column: 2  
Seq primer: T7  
Class: BAC ends.

```

FEATURES
  source
    Location/Qualifiers
      1..462
        /organism="Rattus norvegicus"
        /strain="BN/SSNHsd/MCW"
        /db_xref="taxon:10116"
        /clone="CH230-113M2"
        /clone_lib="CHORI-230 Segment 1"
        /sex="female"
        /cell_type="Brain"
        /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
        CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
        Pieter de Jong"
BASE COUNT      209 a      89 c      87 g      77 t
ORIGIN

```

```

alignment_scores:
  Quality: 70.00      Length: 125
  Ratio: 1.148        Gaps: 4
  Percent Similarity: 48.800      Percent Identity: 24.000

alignment_block:
US-09-528-682-4 x BH287081 ..
Align seg 1/1 to: BH287081 from: 1 to: 462

```

```

3 PheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnTh 19
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 TTCCACCAACAGATAACAGAGATGGAAGAGAGAATCTCAGGACGAGAAGA 101
19 rThrThrGlnPheLeuAsnMetGly..... 27
102 TTCCATAGAAATCATTTGACTCAACTCTCAAGATATGTAAACGGGAAAA 151
28 .....AsnGluValAlaLeuAspGlyArgTyrSer 37
152 AGCTACTGGTCCAAAACATACAGGAATCCAGGAGTCAATGAGAGATCA 201
38 AsnTyrAlaLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheG 54
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
202 AACCTAAGGATAATAGTATAGAGAGAGTGAAGACTCCAGCTCAAAGG 251
54 ySerAlaAsnIle.....AspGlyPhePro..... 62
252 ACCAGTAATATCTTCAACAAATCATAGAGAAAATCTCCCTAACCTAA 301
63 .....GluValArgGluPheAsnSerLeuProAsn... 72
302 AAAAAGAGATACCATAGACATAAAGAGCCTACAGACTCCAAATAGA 351
73 .....AsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAs 87
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 TTGGACCAAGAAAAGAACACCTCCCTCCATACATAATGTCAAAACACCA 401
87 pAlaAspPheLysLysTyrIleLys 95
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 CGCAACAAATAAAGAGATATATAA 426

seq_name: gb_gss:BH260957

```

# seq\_documentation\_block:

LOCUS BH260957 518 bp DNA linear GSS 30-NOV-2001  
DEFINITION CH230-90H4.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone  
CH230-90H4, DNA sequence.  
ACCESSION BH260957  
VERSION BH260957  
KEYWORDS BH260957.1 GI:17165923  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 518)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn,  
A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de  
Jong, P. and Fraser, C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
JOURNAL Unpublished (1999)  
COMMENT Other GSSs: CH230-90H4.TV

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end  
plate: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)  
Plate: 90 row: H column: 4  
Seq primer: SP6  
Class: BAC ends.

## FEATURES

```

  Location/Qualifiers
    1..518
      /organism="Rattus norvegicus"
      /strain="BN/SSNHsd/MCW"
      /db_xref="taxon:10116"
      /clone="CH230-90H4"
      /clone_lib="CHORI-230 Segment 1"
      /sex="female"
      /cell_type="Brain"
      /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
      CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
      Pieter de Jong"
BASE COUNT      232 a      101 c      99 g      86 t
ORIGIN

```

```

alignment_scores:
  Quality: 70.00      Length: 116
  Ratio: 1.186        Gaps: 3
  Percent Similarity: 50.862      Percent Identity: 25.000

alignment_block:
US-09-528-682-4 x BH260957 ..
Align seg 1/1 to: BH260957 from: 1 to: 518

```

```

3 PheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnTh 19
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 TTCCACCAACAGATAACAGAGATGGAAGAGAGAATCTCGGACGAGAAGA 101
19 rThrThrGlnPheLeuAsnMetGly..... 27
102 TTCCATAGATTTTATTGACTCAACTGTCAAGATATGTAAACGGGAAAA 151
28 .....AsnGluValAlaLeuAspGlyArgTyrSer 37
152 AGCTACTGGTCCAAAACATACAGGAATCCAGGACTCAAGGAGAGATCA 201

```

seq\_name: gb\_gss:BH260957















OM of: US-09-528-682-4 to: Issued\_Patents\_NA:\* out\_format : pfs

Date: Jun 18, 2002 7:52 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet-p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09528682/runat\_18062002\_082443\_7827/app\_query.fasta\_1.689  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.rni  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000  
-YGAPEXT=0.500 -GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000  
-XGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human04.cgi -LIST=1000 -DOCALLIGN=200  
-THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682.@CGN1\_1\_96 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-4

Query length: 110

Database: Issued\_Patents\_NA.\*

Database sequences: 383533

Database length: 122816752

Search time (sec): 143.210000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-844-057-1	..	76.00	159.33	0.3466	2619
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-746-111-4	+	76.00	149.08	1.29	6885
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-328-322-3	+	74.50	167.08	0.1283	918
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-328-322-1	+	74.50	163.90	0.1929	1222
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-328-322-6	+	74.50	154.90	0.6120	2745
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-328-322-4	+	74.50	152.20	0.8656	3500
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-212-133A-7	+	71.50	135.91	6.99	7493
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-474-503-5	+	71.50	135.91	6.99	7493
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-670-707A-5	+	71.50	135.91	6.99	7493
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-037-601-5	+	71.50	135.91	6.99	7493
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-037-601-5	+	71.50	135.91	6.99	7493
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-194-290-6	+	71.00	143.73	2.57	3300
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-614-377A-6	+	71.00	143.73	2.57	3300
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-142-648B-6	+	70.50	142.49	3.01	3280
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-729-767-6	+	69.50	158.89	0.3671	594
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-163-833-3	-	69.50	154.12	0.6767	912
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-258-188-1	+	69.00	142.63	2.95	2278
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-526-813-1	+	69.00	142.63	2.95	2278
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-328-322-3	+	69.00	142.63	2.95	2278
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-200-232-3	+	67.00	141.76	3.30	1541
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-200-232-3	+	67.00	141.76	3.30	1541
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-434-000A-5	+	66.50	130.93	13.24	3630
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-312-157-5	+	66.50	130.93	13.24	3630
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-038-682-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-302-833-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-530-198-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-880-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-728-470-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-617-697-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-719-641-3	+	66.00	126.21	24.27	4937
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-038-682-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-302-832-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-469-880-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-880-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-728-470-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-617-697-6	+	66.00	119.14	60.10	9323

/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-719-641-6	+	66.00	119.14	60.10	9323
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-148-680-1	+	65.50	134.99	7.87	1993
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-752-660A-3	+	65.00	136.47	6.51	1552
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-822-322-7	+	64.50	143.16	2.76	756
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-466-109-7	+	64.50	143.16	2.76	756
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-118-469A-4	+	64.50	136.63	6.38	1361
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-056-556-203	+	64.50	136.63	6.38	1361
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-460-739-1	+	64.50	122.43	39.37	4875
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-709-912-19	-	64.00	144.83	2.23	579
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-047-370-19	-	64.00	144.83	2.23	579
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-213-053-1	-	64.00	118.43	65.81	6216
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-765-332-165	+	63.00	150.10	1.13	285
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-448-984-165	+	63.00	150.10	1.13	285
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-044-621D-3	+	63.00	142.35	3.06	572
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-086-633-1	+	63.00	122.06	41.29	3546
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-348-353-16	+	63.00	121.46	44.61	3744
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-963-16	+	63.00	121.46	44.61	3744
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-465-966-16	+	63.00	121.46	44.61	3744
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-478-370-1	+	63.00	114.54	108.33	6973
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-968-685A-9	+	63.00	111.05	169.45	9542
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-330-330-3	+	63.00	110.18	189.55	10322
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-469-667-15	+	62.50	133.57	9.44	1121
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-224-110-15	+	62.50	133.57	9.44	1121
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-224-110-15	+	62.50	133.57	9.44	1121
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-224-110-15	+	62.50	133.57	9.44	1121
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-946-914-1	+	62.50	133.40	9.65	1138
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-302-812-3	+	62.00	117.93	70.19	4069
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-511-477-3	+	62.00	117.93	70.19	4069
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-357-071-1	+	61.50	120.67	49.38	2828
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-038-682-5	+	61.50	107.58	264.46	9171
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-302-832-5	+	61.50	107.58	264.46	9171
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-530-198-5	+	61.50	107.58	264.46	9171
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-728-470-5	+	61.50	107.58	264.46	9171
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-617-697-5	+	61.50	107.58	264.46	9171
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-07-972-791-6	-	61.00	117.44	74.67	3361
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-239-052-1	+	60.50	133.09	10.03	732
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-136-442-2	+	60.50	128.88	17.22	1069
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-858-207A-8	+	60.50	127.63	20.21	1196
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-609-049A-27	+	60.50	108.25	242.73	6831
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-170-996-27	+	60.50	108.25	242.73	6831
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-305-643-3	-	60.00	133.89	9.06	606
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-256	-	60.00	132.13	11.36	710
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-703-947-1	+	60.00	117.39	75.21	2672
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-752-760A-1	+	60.00	88.75	3.0e+03	35081
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-2	+	60.00	35.00	7.4e+05	4037
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-2	+	60.00	34.98	7.4e+05	4037
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-446-777-3	+	59.50	129.85	15.21	775
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-262-220-13	+	59.50	119.13	60.11	2031
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-471-733-13	+	59.50	119.13	60.11	2031
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-468-878-13	+	59.50	119.13	60.11	2031
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-750-494-13	+	59.50	118.90	61.98	2075
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-262-220-3	+	59.50	118.90	61.98	2075
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-471-733-3	+	59.50	118.90	61.98	2075
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-468-878-3	+	59.50	118.90	61.98	2075
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-777-5	+	59.50	117.73	71.96	2304
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-777-5	+	59.50	117.73	71.96	2304
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-714-402-1	+	59.50	112.98	132.30	3531
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-714-402-1	+	59.50	112.98	132.30	3531
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-327-937-1	+	59.50	112.47	141.32	3698
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-470-260-2	+	59.50	111.71	155.81	3960
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-471-491-2	+	59.50	111.71	155.81	3960
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-466-662-2	+	59.50	111.71	155.81	3960
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-466-662-2	+	59.50	111.71	155.81	3960
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-447-411-75	+	59.50	111.22	165.90	4138
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-662-227-33	+	59.50	111.22	165.90	4138
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-017-947-33	+	59.50	111.22	165.90	4138
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-237C-1	+	59.50	85.11	4.7e+03	43280
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-366-13	+	59.00	138.35	5.11	321
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-366-19	+	59.00	138.35	5.11	321
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-366-19	+	59.00	138.35	5.11	321
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-025-769B-281	+	59.00	133.38	9.68	502



















/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-657-392-1 + 51.50 101.99 541.93 1454 ! 1454 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:PCT-US94-02539-1 + 51.50 101.99 541.93 1454 ! 1454 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-657-392-26 + 51.50 101.98 542.46 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:PCT-US94-02539-26 + 51.50 101.98 542.46 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-813-940-1 + 51.50 100.84 628.40 1613 ! 1613 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-133-347-1 + 51.50 100.68 640.66 1635 ! 1635 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-133-347-3 + 51.50 100.68 640.66 1635 ! 1635 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-816-977-46 + 51.50 98.04 898.82 2073 ! 2073 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-352-990-9 + 51.50 97.98 906.25 2085 ! 2085 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-235-838-6 + 51.50 97.22 999.38 2233 ! 2233 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-465-473B-6 + 51.50 97.22 999.38 2233 ! 2233 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-797-266-2 + 51.50 95.86 1.2e+03 2523 ! 2523 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-956-288-2 + 51.50 95.86 1.2e+03 2523 ! 2523 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-07-977-434-11 + 51.50 95.19 1.3e+03 2679 ! 2679 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-458-819-11 + 51.50 95.19 1.3e+03 2679 ! 2679 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:PCT-US91-07035-11 + 51.50 95.19 1.3e+03 2679 ! 2679 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-482-677-5 + 51.50 94.78 1.4e+03 2779 ! 2779 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-152-019A-41 + 51.50 94.77 1.4e+03 2783 ! 2783 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-813-940-3 + 51.50 93.54 1.6e+03 3107 ! 3107 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-157-363A-9 + 51.50 91.23 2.2e+03 3826 ! 3826 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-377-875-1 + 51.50 91.23 2.2e+03 3826 ! 3826 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-922-125-1 + 51.50 91.23 2.2e+03 3826 ! 3826 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:PCT-US93-11405A-9 + 51.50 91.23 2.2e+03 3826 ! 3826 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:PCT-US93-11405-9 + 51.50 91.23 2.2e+03 3826 ! 3826 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-425-843-1 + 51.50 90.57 2.3e+03 4061 ! 4061 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-425-665-1 + 51.50 89.04 2.9e+03 4656 ! 4656 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-685-668-1 + 51.50 89.04 2.9e+03 4656 ! 4656 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-326-117B-1 + 51.50 87.04 3.7e+03 5577 ! 5577 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-982-129-1 + 51.50 87.04 3.7e+03 5577 ! 5577 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-673-312-9 + 51.50 85.93 4.3e+03 6162 ! 6162 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-056-075-1 + 51.50 85.78 4.3e+03 6243 ! 6243 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-103-840A-2 + 51.50 12.84 7.6e+05 4403765 ! 4403765 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-103-840A-1 + 51.50 12.82 7.6e+05 4411529 ! 4411529 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-998-416-906 + 51.00 111.34 163.40 558 ! 558 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-425-315-6 + 51.00 110.35 185.55 610 ! 610 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-716-190-1 + 51.00 110.35 185.55 610 ! 610 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-594-031-89 + 51.00 109.02 219.84 687 ! 687 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-963-743-2 + 51.00 108.27 242.07 735 ! 735 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-463-090B-2 + 51.00 107.53 266.38 786 ! 786 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-094-557-4 + 51.00 104.37 399.35 1044 ! 1044 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-446-803-5 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-861-837-5 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-600-656-5 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-009-494-1 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-170-670-15 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-170-670-15 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-183-412-14 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-183-412-14 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-010-232-3 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-354-191A-5 + 51.00 100.68 641.21 1455 ! 1455 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-232-468A-21 + 51.00 98.94 801.26 1701 ! 1701 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-879-260-3 + 51.00 98.59 838.47 1756 ! 1756 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-231-529-5 + 51.00 98.24 876.18 1811 ! 1811 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-977-816-5 + 51.00 98.24 876.18 1811 ! 1811 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-476-102A-7 + 51.00 98.04 899.73 1845 ! 1845 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-041-981A-5 + 51.00 97.87 918.57 1872 ! 1872 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-247-475-49 + 51.00 97.83 924.18 1880 ! 1880 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-479-650-49 + 51.00 97.83 924.18 1880 ! 1880 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-191-866B-80 + 51.00 97.83 924.18 1880 ! 1880 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-674-169-49 + 51.00 97.83 924.18 1880 ! 1880 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-185-949B-80 + 51.00 97.87 931.20 1890 ! 1890 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-935-855-19 + 51.00 97.21 1.0e+03 1988 ! 1988 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-489-202-26 + 51.00 97.21 1.0e+03 1988 ! 1988 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-484-434C-33 + 51.00 96.97 1.0e+03 2030 ! 2030 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-530-950-1 + 51.00 96.97 1.0e+03 2030 ! 2030 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-149-879-1 + 51.00 96.97 1.0e+03 2030 ! 2030 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-105-390-7 + 51.00 94.17 1.5e+03 2612 ! 2612 !  
/cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-227-689-1 + 51.00 94.07 1.5e+03 2636 ! 2636 !  
/cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-09-368-169-1 + 51.00 94.07 1.5e+03 2636 ! 2636 !







```
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-818-111-11 + 49.50 98.91 804.47 1200 +
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-056-556-11 + 49.50 98.91 804.47 1200 +
/cgn2_6/ptodata/2/ina/backfiles1.seq:5514590-3 + 49.50 97.41 974.86 1373 +
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-931-952-3 + 49.50 96.27 1.1e+03 1521 +
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-272-247-3 + 49.50 96.27 1.1e+03 1521 +
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-08560-3 + 49.50 96.27 1.1e+03 1521 +
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-796-899-22 + 49.50 96.09 1.2e+03 1546 +
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-487-823B-1 + 49.50 96.01 1.2e+03 1558 +
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-997-040-1 + 49.50 96.01 1.2e+03 1558 +
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-03-237-1 + 49.50 96.01 1.2e+03 1558 +
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-472-028A-9 - 49.50 95.05 1.3e+03 1697 +
```

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-844-057-1
```

```
seq_documentation_block:
; Sequence 1, Application US/08844057
; Patent No. 5863777
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5863777el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,057
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607991.8
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmie, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31458-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-844-057-1
```

```
alignment_scores:
Quality: 76.00 Length: 94
Ratio: 1.490 Gaps: 5
Percent Similarity: 54.255 Percent Identity: 30.851
```

```
alignment_block:
US-09-528-682-4 x US-08-844-057-1/rev ..
```

```
Align seg 1/1 to reverse of: US-08-844-057-1 from: 1 to: 2619
```

```
14 lIeAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
||||| :|||:|||||:|||||
2028 GTTCCACATGAGTACACACAAAGTTC.....AATAGAGT 1991
30 lAlaLeuAspGlyArgTyrSerAsn...TyrAlaLeuIleSerAlaGluG 46
```

```
||||| :|||:||||| :|||:||||| :|||:|||||
1990 AGTCACCGATAGTCACGACACAACTGCCTCCGTATTTCACCGAAG 1941
46 lYcGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
||| :|||:|||||:|||||
1940 AGGGCATAGTCCCATTTCTTTAGCAGTGTCAATATC..... 1903
63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
:|||||
1902 .....CGTTTCAAC..... 1894
79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
||||| :|||:||||| :|||:|||||
1893 ...TGCTCTAACTTCACAGAGCTTCCACAGATT.....TTCGTTGACT 1854
96 euleuIleAsnAsnAspGlyPhePheSerAsn 106
|||||:|||||:|||||:|||||
1853 TGCTGTTCAATCGACGCAATTCTTCAGCAGT 1822
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-746-111-4
seq_documentation_block:
; Sequence 4, Application US/08746111
; Patent No. 6066778
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Cui, Jisong
; TITLE OF INVENTION: Compositions And Methods For Screening
; TITLE OF INVENTION: Compounds For Anticoagulant Activity
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,111
; FILING DATE: 06-NOV-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6585 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..6554
US-08-746-111-4
```

```
alignment_scores:
Quality: 76.00 Length: 106
Ratio: 1.169 Gaps: 4
Percent Similarity: 61.321 Percent Identity: 25.472
```

```
alignment_block:
```







```
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,322
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: P38,615
; REFERENCE/DOCKET NUMBER: B600-0151.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2745 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: coding sequence of CNI
; US-08-328-322-6

alignment_scores:
    Quality: 74.50      Length: 95
    Ratio: 1.330        Gaps: 2
Percent Similarity: 58.947 Percent Identity: 26.316

alignment_block:
US-09-528-682-4 x US-08-328-322-6 ..

Align seg 1/1 to: US-08-328-322-6 from: 1 to: 2745

13 ProfileAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGI 29
   ||| ||||:||||: ||| :|||:|: |||
2275 CGAGTGAGAAGCGTTAAATTCACACACTCCTATTTCGCCTCTTAAC.. 2322

29 uValAlaLeuAspGlyArgTyrSerAsnTyrrAlaLeuIleSerAlaGlug 46
   :|||:|||||:||| |::||| ||| :||| :
2323 .ATGAGCTTGTGGAGAGAAAAAATCGTCCAGCTCTACACGCACAT 2371

46 GlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
   :: :: ::
2372 CAGTGTTCCTAGTACAATAAGACAGAACAA..... 2403

63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspH 79
   :|||:||||:||||: ||| :|||:|: |||
2404 .....CAGCATTTCAATGACATAACCAGATGCTTAGCGAGTAGTACGA 2447

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
   :|||:|: ||||:||||: ||| :|||:|: ||| :
2448 AGATGCCHTTCCCAAAGCCAATCATTAATTTCAATAAGAACTACCAG 2497

96 euLeuIleAsnAsnAspGlyPhePheSerSerAsnAsn 107
   :||| |||||:|: |||:|: |||:|: |||
2498 TACTTAAAATTAAATGATAACGTCATACATAACAAAC 2532

seq_name: /cgn2_6/potodata/2/ina/5A_COMB.seq:US-08-328-322-4

seq_documentation_block:
; Sequence 4, Application US/08328322
; Patent No. 5723436
; GENERAL INFORMATION:
; APPLICANT: Huang, Laiqiang
; APPLICANT: Cyert, Martha S.
; TITLE OF INVENTION: Calcineurin Interacting Protein Compos
; TITLE OF INVENTION: and Methods
```







```
; Patent No. 5663060
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 7476
; US-08-212-133A-7

alignment_scores:
  Quality: 71.50      Length: 140
  Ratio: 0.993       Gaps: 5
  Percent Similarity: 51.429  Percent Identity: 21.429

alignment_block:
US-09-528-682-4 x US-08-212-133A-7 ..
Align seg 1/1 to: US-08-212-133A-7 from: 1 to: 7493

4 ThrValArgIleGlnAspAlaTyrGluProIle...AlaAsnThrAsnTh 19
3543 ACAACAGACTTGCAGAGTCCCATATTAAGGTCAATAGTGAGATTCAAGA 3592

19 rThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgT 36
3593 AGTAACAGCTTTGATTTCATGATGGAACACACTTTTAGGCCAAATAATCTCAT 3642

36 yr.....SerAsnTyrAlaLeuIleSerAlaGluGlyMetAspArg 50
3643 ATTCGAGACTAACCATATGCTAAATAGAACTACCTCAACAAAATAATAA 3692

51 AspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPh 67
3693 GACATATTTTCATAGAAAAGATGAAGATCCTATTCCACAAGATGAAGAA 3742

67 eAsnSerLeuProAsnAsn.....PhelystyIle..... 73
3743 TACATCATGCGCATTTTCCAGAGATGTTCTTCTGTCAGAAATCTTCAAAAT 3792

74 .....LysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAla 88
3793 GGTAAAAAGACCAATGGAATAATTCCTGCAACTCGAGCAAGAACAT 3842

89 Asp.....PhelystyIle..... 94
3843 AGTCCAAAGCAATTAGTATATTATGTTTAAATAATATGTAATAATCA 3892

95 .....LysLeuLeuIleAsnAsnAspGlyPheP 104
3893 AAGTTTCTTGTCAGAGAAAAATAAAGTCACAGTACAGACAGGATGGATTTA 3942

104 heSerAsnAspGlyLys 110
3943 CAAGAACAATAGGACTTAAA 3962

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-474-503-5
seq_documentation_block:
; Sequence 5, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: repeat_unit
; LOCATION: 1..407
; OTHER INFORMATION: /rpt_type="terminal"
; OTHER INFORMATION: /note="5'UTR"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7471..7476
; OTHER INFORMATION: /function="PolyA_signal"
; FEATURE:
; NAME/KEY: repeat_unit
; LOCATION: 7368..7493
; OTHER INFORMATION: /rpt_type="terminal"
; OTHER INFORMATION: /note="3'UTR"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 408..7367
; OTHER INFORMATION: /product="Coagulation Factor VIII"
; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Lakich, D.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the Murine Factor VIII cDNA.
; Patent No. 5744446
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 7476
; US-08-474-503-5
```

```
alignment_scores:
  Quality: 71.50      Length: 140
  Ratio: 0.993       Gaps: 5
  Percent Similarity: 51.429  Percent Identity: 21.429

alignment_block:
US-09-528-682-4 x US-08-474-503-5 ..
Align seg 1/1 to: US-08-474-503-5 from: 1 to: 7493

4 ThrValArgIleGlnAspAlaTyrGluProIle...AlaAsnThrAsnTh 19
3543 ACAACAGACTTGCAGAGTCCCATATTAAGGTCAATAGTGAGATTCAAGA 3592

19 rThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgT 36
3593 AGTAACAGCTTTGATTTCATGATGGAACACACTTTTAGGCCAAATAATCTCAT 3642

36 yr.....SerAsnTyrAlaLeuIleSerAlaGluGlyMetAspArg 50
3643 ATTCGAGACTAACCATATGCTAAATAGAACTACCTCAACAAAATAATAA 3692

51 AspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPh 67
3693 GACATATTTTCATAGAAAAGATGAAGATCCTATTCCACAAGATGAAGAA 3742

67 eAsnSerLeuProAsnAsn.....PhelystyIle..... 73
3743 TACATCATGCGCATTTTCCAGAGATGTTCTTCTGTCAGAAATCTTCAAAAT 3792

74 .....LysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAla 88
3793 GGTAAAAAGACCAATGGAATAATTCCTGCAACTCGAGCAAGAACAT 3842

89 Asp.....PhelystyIle..... 94
3843 AGTCCAAAGCAATTAGTATATTATGTTTAAATAATATGTAATAATCA 3892

95 .....LysLeuLeuIleAsnAsnAspGlyPheP 104
3893 AAGTTTCTTGTCAGAGAAAAATAAAGTCACAGTACAGACAGGATGGATTTA 3942

104 heSerAsnAspGlyLys 110
3943 CAAGAACAATAGGACTTAAA 3962
```



```
3593 AGTAACAGCTTTGATTTCATGATGAACACTTTTAGGCAAAAAATCTACAT 3642
36 yr.....SerAsnTyrAlaLeuIleSerAlaGluGlyMetAspArg 50
||| : : : : :
3643 ATTGAGACTAAACCATATGCTTAATAGAACACTACCTCAACAAAAATAAA 3692
||| : : : : :
51 AspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPh 67
||| : : : : :
3693 GACATATTTTCATAGAAAGATGAAGATCCTATTCACAGATGAAGAGAA 3742
67 eAsnSerLeuProAsnAsn..... 73
: : : : :
3743 TACAATCATGCCATTTTCCAGAGATGTTGTTCTGTGAGAAATCTCAAAAT 3792
74 .....LysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAla 88
||| : : : : :
3793 GGTTTAAAGAACCAATGGAATAATTCCTTGAACCTGAGCAAGAACAT 3842
89 Asp.....PheLysLysTyrIle..... 94
: : : : :
3843 AGTCCAAAGCAATTAGTATATTATTTAATGTTTAAAAAATATGTAATAATCA 3892
95 .....LysLeuLeuIleAsnAsnAspGlyPheP 104
||| : : : : :
3893 AAGTTTCTTGTCCAGAGAAAAATAAAGTCACAGTAGAACAGGATGGATTTA 3942
104 heSerAsnAsnGlyLys 110
: : : : :
3943 CAAGAACAATAGGACTTAA 3962

seq_name: /cgn2_6/ptodata/2/ins/5B_COMB.seq:US-08-670-707A-5
seq_documentation_block:
; Sequence 5, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 7493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: repeat_unit
; LOCATION: 1..407
; OTHER INFORMATION: /rpt_type= "terminal"
; OTHER INFORMATION: /note= "5' UTR"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7471..7476
; OTHER INFORMATION: /function= "polyA signal"
; FEATURE:
; NAME/KEY: repeat_unit
; LOCATION: 7388..7493
; OTHER INFORMATION: /rpt_type= "terminal"
; OTHER INFORMATION: /note= "3' UTR"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 408..7367
; OTHER INFORMATION: /product= "coagulation factor VIII"
; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Lakich, D.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the murine Factor VIII cDNA
; Patent No. 5859204
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; US-08-670-707A-5

alignment_scores:
Quality: 71.50 Length: 140
Ratio: 0.993 Gaps: 5
Percent Similarity: 51.429 Percent Identity: 21.429

alignment_block:
US-09-528-682-4 x US-08-670-707A-5 ..
Align seg 1/1 to: US-08-670-707A-5 from: 1 to: 7493
4 ThrValArgIleGlnAspAlaTyrGluProIle...AlaAsnThrAsnTh 19
||| : : : : :
3543 ACAACAGACTTGCAAGATGCCATATTAAGGTCAATAGTGAGATTCAAGA 3592
19 rThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgT 36
: : : : :
3593 AGTAACAGCTTTGATTTCATGATGAACACTTTTAGGCAAAAAATCTACAT 3642
36 yr.....SerAsnTyrAlaLeuIleSerAlaGluGlyMetAspArg 50
||| : : : : :
3643 ATTGAGACTAAACCATATGCTTAATAGAACACTACCTCAACAAAAATAAA 3692
51 AspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPh 67
||| : : : : :
3693 GACATATTTTCATAGAAAGATGAAGATCCTATTCACAGATGAAGAGAA 3742
67 eAsnSerLeuProAsnAsn..... 73
: : : : :
3743 TACAATCATGCCATTTTCCAGAGATGTTGTTCTGTGAGAAATCTCAAAAT 3792
74 .....LysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAspAla 88
||| : : : : :
3793 GGTTTAAAGAACCAATGGAATAATTCCTTGAACCTGAGCAAGAACAT 3842
89 Asp.....PheLysLysTyrIle..... 94
: : : : :
3843 AGTCCAAAGCAATTAGTATATTATTTAATGTTTAAAAAATATGTAATAATCA 3892
95 .....LysLeuLeuIleAsnAsnAspGlyPheP 104
||| : : : : :
3893 AAGTTTCTTGTCCAGAGAAAAATAAAGTCACAGTAGAACAGGATGGATTTA 3942
104 heSerAsnAsnGlyLys 110
: : : : :
3943 CAAGAACAATAGGACTTAA 3962
```











```

; TELEPHONE: 703-684-5600
; TELEFAX: 703-836-5288
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Caulobacter crescentus
; STRAIN: CB 15
; US-08-194-290-6

alignment_scores:
    Quality: 71.00      Length: 112
    Ratio: 1.164       Gaps: 5
    Percent Similarity: 54.464      Percent Identity: 28.571

alignment_block:
US-09-528-682-4 x US-08-194-290-6 ..

Align seg 1/1 to: US-08-194-290-6 from: 1 to: 3300

1 AsptyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 GACTACCTGACCGCCTTCGTGGCG.....GCCAACAC 639

17 rAsn...ThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
640 GCCGTTCCAGCGCGCTGCCGACATCGATCTGGCGCTCAAGCGCGCTGA 689

33 spGlyArgTyrSerAspTyrAlaLeuIleSerAlaGluGly..... 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
690 TCGGCACCATCTGAAACCGCGCCACCGTGTCTGGCGATCGTGTACGCG 739

48 .....MetAspArgAspLeuPhe...GlySerAlaAsnIle 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
740 ACCGCCACGCGCGCGATGATCAACGACCTGTCCGACGCGCCCTGTCCG 789

58 eAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLys 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
790 CGACACGCGCGCTGGCGTGAACCTGTTCACCGCTATCGTGTGGCGG 839

75 laSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLys 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
840 TGTCGGTTTCGACCTCTCGCTGACCAACCGCGCACACACC..... 880

92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103
|||:|||||:|||||:|||||:|||||:|||||:
881 .....CTGACGGGACCGGCCAACACGACACGTTTC 910

seq_name: /cqn_2.6/ptodata/2/ina/5B_COMB.seq:US-08-614-377A-6

seq_documentation_block:
; Sequence 6, Application US/08614377A
; Patent No. 5976864
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingler, Wade H.
; APPLICANT: No. 5976864ellini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF
; TITLE OF INVENTION: HETEROLOGOUS
; TITLE OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson PC
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/614.377A
; FILING DATE: 12-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/194,290
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/895,367
; FILING DATE: 09-JUNE-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsao, Y. Rocky
; REGISTRATION NUMBER: 34053
; REFERENCE/DOCKET NUMBER: 08106/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Caulobacter crescentus
; STRAIN: CB 15
; US-08-614-377A-6

alignment_scores:
    Quality: 71.00      Length: 112
    Ratio: 1.164       Gaps: 5
    Percent Similarity: 54.464      Percent Identity: 28.571

alignment_block:
US-09-528-682-4 x US-08-614-377A-6 ..

Align seg 1/1 to: US-08-614-377A-6 from: 1 to: 3300

1 AsptyrPheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 GACTACCTGACCGCCTTCGTGGCG.....GCCAACAC 639

17 rAsn...ThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
640 GCCGTTCCAGCGCGCTGCCGACATCGATCTGGCGCTCAAGCGCGCTGA 689

33 spGlyArgTyrSerAspTyrAlaLeuIleSerAlaGluGly..... 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
690 TCGGCACCATCTGAAACCGCGCCACCGTGTCTGGCGATCGTGTACGCG 739

48 .....MetAspArgAspLeuPhe...GlySerAlaAsnIle 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
740 ACCGCCACGCGCGCGATGATCAACGACCTGTCCGACGCGCCCTGTCCG 789

58 eAspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLys 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
790 CGACACGCGCGCTGGCGTGAACCTGTTCACCGCTATCGTGTGGCGG 839

75 laSerSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLys 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
840 TGTCGGTTTCGACCTCTCGCTGACCAACCGCGCACACACC..... 880

92 LysTyrIleLysLeuLeuIleAsnAsnAspGlyPhe 103
|||:|||||:|||||:|||||:|||||:|||||:
881 .....CTGACGGGACCGGCCAACACGACACGTTTC 910

```























```

; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02219
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 2200.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1541
; PCT-US95-02219-3

alignment_scores:
  Quality: 67.00      Length: 96
  Ratio: 1.117      Gaps: 2
  Percent Similarity: 62.500      Percent Identity: 22.917

alignment_block:
US-09-528-682-4 x PCT-US95-02219-3

Align seg 1/1 to: PCT-US95-02219-3 from: 1 to: 1541

11 TyrGluProIleAlaAsnThrAsnThrThrThrGlnPheLeuAsnMetG1 27
1210 TACAACCACTCATTAAGATCAATAACGCTCAAAATCTCACTAAATAA 1259
27 yAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSera 44
1260 AGAACATGTTTAGTGAAGGCGCAACATTGATTATAATTAGTGGGAG 1309
44 laGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGly 60
1310 TGCAAGGCGCT...AGTTATGACAATATTTCTGCAAGCAACCAATCTG 1356
61 PheProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSe 77
1357 CAAGACGAATTCAAGAGCGCTAGCCCTTTATACACACACACACCGCAT 1406
77 rAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrI 94
1407 GGATATTTGTGTGGTGGCGAAGGAACGATACCGATGACATTAAACA...T 1453
94 leLysLeuLeuIleAsnAsnAspGlyPhePheSerAsn 106
1454 GCGGGATGGCTATCGGCAATCAAGCATGGTGAATAAC 1491

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-02219A-3

```

```

seq_documentation_block:
; Sequence 3, Application PC/TUS9502219A
; GENERAL INFORMATION:
; APPLICANT: Cover, Timothy L.
; APPLICANT: Tummuru, Murali KR
; APPLICANT: Cao, Ping
; APPLICANT: Thompson, Stuart A.
; APPLICANT: Blaser, Martin J.
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02219A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 2200.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1541
; PCT-US95-02219A-3

alignment_scores:
  Quality: 67.00      Length: 96
  Ratio: 1.117      Gaps: 2
  Percent Similarity: 62.500      Percent Identity: 22.917

alignment_block:
US-09-528-682-4 x PCT-US95-02219A-3

Align seg 1/1 to: PCT-US95-02219A-3 from: 1 to: 1541

11 TyrGluProIleAlaAsnThrAsnThrThrThrGlnPheLeuAsnMetG1 27
1210 TACAACCACTCATTAAGATCAATAACGCTCAAAATCTCACTAAATAA 1259
27 yAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSera 44
1260 AGAACATGTTTAGTGAAGGCGCAACATTGATTATAATTAGTGGGAG 1309
44 laGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGly 60
1310 TGCAAGGCGCT...AGTTATGACAATATTTCTGCAAGCAACCAATCTG 1356
61 PheProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSe 77
1357 CAAGACGAATTCAAGAGCGCTAGCCCTTTATACACACACACACCGCAT 1406

```















```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-880-3

alignment_scores:
  Quality: 66.00      Length: 95
  Ratio: 1.179      Gaps: 3
  Percent Similarity: 58.947      Percent Identity: 25.263

alignment_block:
US-09-528-682-4 x US-08-469-880-3 ..

Align seg 1/1 to: US-08-469-880-3 from: 1 to: 4937

15 AlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValal 31
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3439 GCCAACCTAACTATTAAACCAAGAAATTGAAATTCACAGAACCTAAG 3488

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyLys 48
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3489 TATTTTCAGGT.....TTCAATAAGCAGAGATTACAGCCAAAGATGGTA 3532

48 etAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3533 GAGATTTAACCTATTGGCAACAGTAAATGACGTAACAGCGTGCCGAAGCC 3582

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3583 AAACAGTAACCTTTTAAACAATGTAAAGATTCAAAAATCTCTGCTGAC.. 3630

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3631 .....GGTCACAATGTGACACTAAATAGCAAAAGTGAAAA 3664

96 euLeuIleAsnAspGlyPhePheSerAsnAsn 107
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3665 CATCTAGCAGCAATGCGGACGTAACAGCAATAGC 3699

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-728-470-3

seq_documentation_block:
; Sequence 3, Application US/08728470
; Patent No. 5928651
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,880
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-516 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```



```
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-728-470-3

alignment_scores:
    Quality: 66.00      Length: 95
    Ratio: 1.179       Gaps: 3
Percent Similarity: 58.947 Percent Identity: 25.263

alignment_block:
US-09-528-682-4 x US-08-728-470-3 ..

Align seg 1/1 to: US-08-728-470-3 from: 1 to: 4937

15 AlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3439 GCCAACCTTAACCTATTAACCAAGAAATGGAATTGACAGAACCTAAG 3488

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyM 48
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3489 TATTTTCAGGT.....TTCATAAAGCAGAGATTACAGCCAAAGATGGTA 3532

48 etAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3533 GAGATTTAACTATTGGCAACAGTAATGACGGTAACAGCGGTGCCGAAGCC 3582

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3583 AAAACAGTAACCTTTTAAACATGTTAAAGATTCAAAAATCTCTGCTGAC.. 3630

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3631 .....GGTCACAAATGTGACACTTAAATAGCAAAAGTGAATA 3664

96 euLeuIleAsnAsnAspGlyPhePheSerAsnAsn 107
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3665 CATCTAGCAGCAATGGCGGACGTGAAGCAATAGC 3699

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-617-697-3

seq_documentation_block:
; Sequence 3, Application US/08617697
; Patent No. 5977336
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,697
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 05-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-617-697-3

alignment_scores:
    Quality: 66.00      Length: 95
    Ratio: 1.179       Gaps: 3
Percent Similarity: 58.947 Percent Identity: 25.263

alignment_block:
US-09-528-682-4 x US-08-617-697-3 ..

Align seg 1/1 to: US-08-617-697-3 from: 1 to: 4937

15 AlaAsnThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3439 GCCAACCTTAACCTATTAACCAAGAAATGGAATTGACAGAACCTAAG 3488

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyM 48
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3489 TATTTTCAGGT.....TTCATAAAGCAGAGATTACAGCCAAAGATGGTA 3532

48 etAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3533 GAGATTTAACTATTGGCAACAGTAATGACGGTAACAGCGGTGCCGAAGCC 3582

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3583 AAAACAGTAACCTTTTAAACATGTTAAAGATTCAAAAATCTCTGCTGAC.. 3630

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3631 .....GGTCACAAATGTGACACTTAAATAGCAAAAGTGAATA 3664

96 euLeuIleAsnAsnAspGlyPhePheSerAsnAsn 107
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3665 CATCTAGCAGCAATGGCGGACGTGAAGCAATAGC 3699

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-08-719-641-3

seq_documentation_block:
; Sequence 3, Application US/08719641
; Patent No. 6218141
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

alignment_block:
US-09-528-682-4 x US-08-719-641-3 ..
Align seg 1/1 to: US-08-719-641-3 from: 1 to: 4937

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetClyAsnGluValAl 31
||||| : : : : : ||| : : : : :
3439 GCCAACCTTAACTATTAAACCAAGAAATTTGAAATTGACAGACACCTAAG 3488
||||| : : : : : ||| : : : : :
31 aleuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyLys 48
: : : : : ||| : : : : : ||| : : : : :
3489 TATTTCAGGT.....TTCAATAAAGCAGAGATTACAGCCAAAGATGGTA 3532
||||| : : : : : ||| : : : : :
48 etAspArgAspLeuPheGlySerAlaAsnIleAspClyPheProGluVal 64
||||| : : : : : ||| : : : : :
3533 GAGATTTAACATTATGGCAACAGCTAATGACGGTAAACAGCGTGCAGAGCC 3582
||||| : : : : : ||| : : : : :
65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : ||| : : : : : ||| : : : : :
3583 AAAACAGTAACATTTAACAATGTTAAAGATTCAAAAATCTCTGCTGAC... 3630
||||| : : : : : ||| : : : : :
79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
||||| : : : : : ||| : : : : :
3631 .....GPTCACAATGTGACACTAAATAGCAAGTCAAAA 3664

```

```

alignment_scores:      Quality: 68.00      Length: 95
                       Ratio:   1.179     Gaps:    3
Percent Similarity: 58.947 Percent Identity: 25.263

alignment_block:
US-09-528-682-4 x US-08-038-682-6 ..

Align seg 1/1 to: US-08-038-682-6 from: 1 to: 9323

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31
||||| ::: ||||::: ||||::: ::::||:::

3879 GCCACCTAACATATTAAACCAAGAAAGTAATTGAATTCAGACAGACCTAAG 3928

31 aLeuAspGlyArgTyrSerAsnTyralaLeuIleSerAlaGluGlyGlyM 48
:::::|||| |::| ||| ::|||::: |||
3929 TATTTTCAGGT.....TTCAATAAAGCAGAGATTACAGCCAAAAGATGGTA 3972

48 eTasPArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
|||| ::::||||| ::::||||| ::::|||||
3973 GAGATTTAACACTATTGCCAACACGATATGACGGTATNACAGCGGTGCCGAAGCC 4022

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79

```







[illegible]



```

; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-728-470-6

alignment_scores
Quality: 66.00 Length: 95
Ratio: 1.179 Gaps: 3
Percent Similarity: 58.947 Percent Identity: 25.263

alignment_block
US-09-528-682-4 x US-08-728-470-6 ..

Align seg 1/1 to: US-08-728-470-6 from: 1 to: 9323

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAl 31
||||| ::| ||||| ::| ||||| ::| ||||| ::| ||||| ::|
3879 GCCACCTTAACATTAAACCAAGAAATTGAATTCACAGAACCTAAG 3928

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyM 48
:::::|||| ||| ||| |||||::::: |||
3929 TATTTTCAGT.....TTCAATAAAGCAGACGATTACAGCCAAAGATGGTA 3972

48 eTAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||| ::| ||||| ::| ||||| ::| ||||| ::| ||||| ::|
3973 GAGATTTAACATTATGGCAACAGTAATGACGGTAACAGCGGTGCCGAAGCC 4022

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
::| |||||::::: ::| ||||| ::| ||||| ::| ||||| ::|
4023 AAACAGTAACATTTAACAATGTTAAAGATTCAAAAATCTGCTGAC... 4070

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysTyrIleLysL 96
||||| ::| ||||| ::| ||||| ::| ||||| ::| ||||| ::|
4071 .....GGTCACAAATGTGACATAAATAGCAAAAGTGA AAA 4104

96 eUleuIleAsnAsnAspGlyPhePheSerSerAsnAsn 107
::| ||| ||| ||||| ::|
4105 CATCTAGCAGCAATGCGGACGCTGAAAGCAATAGC 4139

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-617-697-6

seq_documentation_block:
; Sequence 6, Application US/08617697
; Patent No. 5977336
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.

```

```
; ZIP: 2202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/617,697
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 05-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-617-697-6

alignment_scores:
      Quality:    66.00       Length:   95
          Ratio:    1.179        Gaps:     3
Percent Similarity: 58.947 Percent Identity: 25.263

alignment_block:
US-09-528-682-4 x US-08-617-697-6 ..

Align seg 1/1 to: US-08-617-697-6 from: 1 to: 9323

15 AlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValal 31
||||| :: ||||| ::||::| :|::|::| |::|::|::|
3879 GCCAACCTAACTATTAAACCAAGGAATTTGAAATTCACAGAAGCCTAAG 3928

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeulleSerAlaGluGlyLys 48
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3929 TATTTCCAGGT.....TTCAAATAAAGCAGAGATTACAGGCCAAAAGATGGTA 3972

48 eTASPARGasPLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||||| :::|::|::|::|::|::|::|::|::|::|::|::|
3973 GAGATTTAACTATTGGCAACAGTAATGACCGGTAAACAGCGTGCGCAAGCC 4022

65 Arg.....GLUPHeASnSerLeuProAsnAsnLysAlaSerSerAspTh 79
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
4023 AAAACAGTAACTTTTACAACATGTTAAGATTCCAATACTCTCTGCAC... 4070

79 rLaSerLeuAsnLysGlnHisAspAlaaspPheLysLysTyrIleLysL 96
|||::|::|::|::|::|::|::|::|::|::|::|::|::|
4071 .....GGTCACAAATGTGACACTAAATAGCAAAAGTAAAA 4104

96 euLeuIleAsnAsnAspGlyPhePheSerSerAsn 107
::|::|::|::|::|::|::|::|::|::|::|::|::|
4105 CATCTAGCACGACCATGCCGACGTGAAAGCAANTAGC 4139

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-719-641-6

seq_documentation_block:
; Sequence 6, Application US/08719641
; Patent No. 6218141
; GENERAL INFORMATION:
```



```

alignment_scores:
  Quality: 66.00      Length: 95
  Ratio: 1.179      Gaps: 3
  Percent Similarity: 58.947      Percent Identity: 25.263

alignment_block:
  US-09-528-682-4 x US-08-719-641-6  ..

Align seg 1/1 to: US-08-719-641-6 from: 1 to: 9323

15 AlaAsnThrAsnThrThrThrClnPheLeuAsnMetGlyAsnGluValAl 31
||||| : : : : : ||| : : : : :
3879 GCCACCTTAACATTAAACACCAAGAATTTGAAATTGACAGACACCTAAG 3928

31 aLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyLys 48
: : : : : ||| ||| ||| ||| : : : : : |||
3929 TATTTCAGGT.....TTCAATTAAGACGAGATTTACGCGCAAGATGGTA 3972

48 eTAspArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluVal 64
||| : : : : : ||| : : : : :
3973 GAGATTTAACATTATGGCAACACATAATGACGGTGAACAGCGGTGCCGAGCC 4022

65 Arg.....GluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : ||| ||| ||| : : : : : ||| : : : : :
4023 AAAACAGTAACCTTTAAACAACTCTTAAAGATTTCAAAATCTCTGCTGAC.. 4070

```

```
alignment_scores:
  Quality: 65.50      Length: 124
  Ratio: 1.129       Gaps: 6
  Percent Similarity: 46.774    Percent Identity: 26.613

alignment_block:
```



```

;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1552 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-752-760A-3

alignment_scores:
Quality: 65.00 Length: 77
Ratio: 1.275 Gaps: 3
Percent Similarity: 66.234 Percent Identity: 24.675

alignment_block:
US-09-528-682-4 x US-08-752-760A-3 ..

Align seg 1/1 to: US-08-752-760A-3 from: 1 to: 1552

27 GlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSe 43
|||||:|||||: : : : : |||:|||||:
69 GGAGGTTCGGTTTGTCCCTCGCGGTATATGGCTCCTACGGA..... 110
||| : : : : : |||:|||||:
43 rAlaGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspG 60
||| : : : : : |||:|||||:
111 .....GGCGAGAAACAGCATTCGTTACTCGGAGCTGGCTCCGTT... 149
60 lypHeProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSer 76
|||||: : : : : |||:|||||:
150 .....GTACGACACACCTCGCGTGTACTTGGTGACAAACAAGTCG 188
77 SerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTy 93
: : : : : |||:|||||:
189 CGGACATCGCTTCCTGNACTATCAAAAC.....GACCACAGCAACTT 232
93 rIleLysLeuLeuIleAsnAsnAspGlyPhe 103
: : : : : |||:|||||:
233 CCTGACCACCGTGGTGACAGAAACAAGATTTC 263

seq_name: /cgn2_5/ptodata/2/ina/8A_COMB.seq:US-08-822-322-7

seq_documentation_block:
; Sequence 7, Application US/08822322
; Patent No. 6037158
; GENERAL INFORMATION:
; APPLICANT: Hummel, Werner, Riebel, Bettina
; TITLE OF INVENTION: Alcohol dehydrogenase and its use for
; TITLE OF INVENTION: enzymatic production of chiral hydroxy compounds
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,322
; FILING DATE: 21-March-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 10 984
; FILING DATE: 21-March-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6037158man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1076

```







```

; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,469A
; FILING DATE: 08-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,757
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: YU-102CIP
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1361 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 109..1359
;
US-08-118-469A-4

alignment_scores:
  Quality: 64.50      Length: 98
  Ratio: 1.316      Gaps: 3
  Percent Similarity: 50.000      Percent Identity: 23.469

alignment_block:
US-09-528-682-4 x US-08-118-469A-4 ..
Align seg 1/1 to: US-08-118-469A-4 from: 1 to: 1361

24 LeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyr.. 39
||||| |||||
244 CTAATTTCTAAGAACAATAGGCTTAAGATTCTCGGTTAAGTAATTTGA 293
39 .....

294 AGCAAAAAAATGACGACATTAATAAATCCAAAGACTTTAAAAAGG 343
40 .....AlaLeuIleSerAlaGlyGlyMetAspArgAspLeuPhe 53
||||| |||||
344 ATTTACAAACTTTAAGAAATTCAAAATTTAATGCTTAAGACTTTGGAT 393
54 GlySerAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLe 70
||||| |||||
394 CAGTCGAGTAAT .....GATTTTCAAAATTT 419
70 uProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHis 87
||||| |||||
420 AGACAATTTCTGAGCTTTCGAAAGAGCTTCTTCA .....AAGCACA 460
87 spAlaAspPheLysLysTyrIleLysLeuLeuIleAsnAsnAsp 101
```

```

;
; 461 ATATTGGCAAGTCAAGATACGGTAAAGCTTTGCTGAAAAATGAT 504
;
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-909-119-4

seq_documentation_block:
; Sequence 4, Application US/08909119
; Patent No. 5807685
; GENERAL INFORMATION:
; APPLICANT: Flavell, Richard A.
; APPLICANT: Fikrig, Erol
; APPLICANT: Lam, Tuan T.
; APPLICANT: Kantor, Fred S.
; APPLICANT: Barthold, Stephen W.
; TITLE OF INVENTION: NOVEL B. BURGENDORFERI POLYPEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/909,119
; FILING DATE: 11-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/118,469
; FILING DATE: 08-SEP-1993
; APPLICATION NUMBER: US 08/099,757
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: YU-102CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1361 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 109..1359
;
US-08-909-119-4

alignment_scores:
  Quality: 64.50      Length: 98
  Ratio: 1.316      Gaps: 3
  Percent Similarity: 50.000      Percent Identity: 23.469

alignment_block:
US-09-528-682-4 x US-08-909-119-4 ..
Align seg 1/1 to: US-08-909-119-4 from: 1 to: 1361

24 LeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyr.. 39
||||| |||||
244 CTAATTTCTAAGAACAATAGGCTTAAGATTCTCGGTTAAGTAATTTGA 293
```



```
39 ..... 39
294 AAGCAAAAAAATGACCAGACATTAAAAAATCCAAAGACTTTAAAAAGG 343
40 .....AlaLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPhe 53
344 ATTTACAAACTTTAAGAAATTAAGAAATTAATGCTTAAGACTTGGAT 393
54 GlySerAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLe 70
394 CAGTCGAGTAAT.....GATTITGAAATTT 419
70 uProAsnAsnLysAlaSerAspThrAlaSerLeuAsnLysGlnHisA 87
420 AGACAATTCTGAGCTTTGCAAGAAGCTTCTTCA.....AAGCACA 460
87 spAlaAspPheLysLysTyrIleLysLeuLeuLeuAsnAsnAsp 101
461 ATATTGGCAAGTCAAGATACGGTAAAGCTTTGCTGAAAAATGAT 504
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-056-556-203
```

```
seq_documentation_block:
; Sequence 203, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; TREATMENT OF HIV INFECTION
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 203:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-056-556-203
```

alignment\_scores:  
Quality: 64.50 Length: 126  
Ratio: 1.057 Gaps: 8  
Percent Similarity: 48.413 Percent Identity: 26.984

alignment\_block:  
US-09-528-682-4 x US-09-056-556-203 ..  
Align seg 1/1 to: US-09-056-556-203 from: 1 to: 2852

```
10 AlaTyrGluProIleAlaAsnThrAsnThr..... 19
724 AGCTACAAACCG...GGCACAGCAATACCGCGGCTTCAACATGGGCCA 770
20 ....ThrThrGlnPheLeuAsnMetGlyAsn.....GluValAlaL 32
771 GTCAACACACGGCTACTGTAACACGCGCACTACAACACCGCGCTTGCCA 820
32 euAspGlyArgTyrSerAsnTyrAlaLeuIleSerAla..... 44
821 ACTCCGGCAATGCAACACCGCGGCTTCAATATTGCGCAACTTCAACAAC 870
45 .....GluGlyGlyMetAspArgAspLeuPheGlySerAlaAs 57
871 GGCTTTCTGTGGCGCGGCGACCAAGGCTGATTTTCGGGAGC..... 915
57 nIleAspGlyPheProGluValArgGluPheAsnSerLeuPro..... 71
916 ....CCCGGCTTC.....TTCAACTCGACCAAGTGCCTGCGCGT 946
72 .....AsnAsnLysAlaSerAspThrAlaSerLeuAsn 83
947 CGTCGGGATTCTTCAACACGCGGTAGCGCGCTCCGCTTCTCTGAAC 996
84 LysGlnHisAspAlaAspPheLysLysTyrIleLysLeuLeuLeuAsn 100
997 TCCGGT.....GCCAACAA 1010
100 nAspGlyPheSerAsnAsnGlyGly 109
1011 TTCTGGCTTCTTCAACTCTTCGTCGGGG 1038
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-460-739-1
seq_documentation_block:
; Sequence 1, Application US/08460739
; Patent No. 5698416
; GENERAL INFORMATION:
; APPLICANT: Wolf, Marcia K.
; APPLICANT: Cassels, Frederick J.
; APPLICANT: Bell, Brian A.
; TITLE OF INVENTION: Improved Methods for Production of
; TITLES OF INVENTION: Antigens Under Control of Temperature-Regulated Promoters
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glenna Hendricks
; STREET: 9669 A Main Street
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,739
; FILING DATE: 20-APR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna M.
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: wolf2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
```



```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: DNA for CS6 protein
; US-08-460-739-1

alignment_scores:
  Quality: 64.50      Length: 101
  Ratio: 1.240        Gaps: 2
Percent Similarity: 51.485 Percent Identity: 20.792

alignment_block:
US-09-528-682-4 x US-08-460-739-1 ..
Align seg 1/1 to: US-08-460-739-1 from: 1 to: 4875

1 AsptYrPhrThrValArgIleGlnAspAlaTyrGluProIleAlaAsnTh 17
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
2831 AATTATTATTTCATGCGCGTGGCGTAATCTTTTGAATTTAAACCAAAACAA 2880
17 rAsnThrThrThrGlnPheLeuAsn..... 25
  ::::: |||
2881 CGATGAAAAAACATCGGAAAGAACTACACTATCTAGAAAAAAGCTTTT 2930
26 .....MetGlyAsnGluValAlaLeuAspGly 34
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
2931 ATGATAAAAGCTAAACTAGTCGTGGTGAAGTTATACGAATTCAAAT 2980
35 ArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAspArgAs 51
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
2981 GTTTATAATAACTACTCTTTTACTGCTGTTTTCAGTTTCTACAGATACAGA 3030
51 pleuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheA 68
  ::::::: :::::::::::::: :::::::::::::: ::::::::::::::
3031 TATGTATACGCCAAGTGAATCGATTATACACACAGAAAT..... 3070
68 snSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLys 84
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
3071 .....CATGGAGTGGCTGATTACAGACTCTCAGATTATTCTCAGG 3109
85 Gln 85
  |||
3110 CAA 3112

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-709-912-19

seq_documentation_block:
; Sequence 19, Application US/08709912
; Patent No. 5759840
; GENERAL INFORMATION:
; APPLICANT: Sung Dr., Wing L
; APPLICANT: Yaguchi Dr., Makoto
; APPLICANT: Ishikawa Dr., Kazuhiko
; TITLE OF INVENTION: Modification of xylanase to improve
; TITLE OF INVENTION: Thermophilicity, Alkalophilicity and
; TITLE OF INVENTION: Thermostability
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitzpatrick, Cella, Harper, and Scinto
; STREET: 277 Park Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10172-0194
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,912
```

```
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Olsen Mr, Warren E
; REGISTRATION NUMBER: 27290
; REFERENCE/DOCKET NUMBER: 1039.2000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-2400
; TELEFAX: (212) 758-2982
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: pYbc
; POSITION IN GENOME:
; UNITS: bp
; US-08-709-912-19

alignment_scores:
  Quality: 64.00      Length: 92
  Ratio: 1.103        Gaps: 5
Percent Similarity: 63.043 Percent Identity: 31.522

alignment_block:
US-09-528-682-4 x US-08-709-912-19/rev ..
Align seg 1/1 to reverse of: US-08-709-912-19 from: 1 to: 579

5 ValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsn..ThrThr 20
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
440 GTGACGTAATGTTGGCGTTCGAACACGTCGCGCGCTTAGATTGCTAAC 391
21 ThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSe 37
  ||| :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
390 ACTCCAATACTAGTAAAGGTGGTCGATCGCATCGATGGAAGG..... 346
37 rAsnTyrAlaLeuIleSerAlaGluGly..... 46
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
345 .....TCGGTTGTATCTTTGGTGGTGTAGATGTCATAGGTACCACCA 303
47 ..GlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGly...Phe 61
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
302 TCGCTTTTGACTGTGCCCTTTGTAGGTTCCAGTCG....GACGGTAGGTTTC 257
62 ProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAs 78
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
256 CCAAGAGTCGACGACGTAATATTCAATCAGTGGCGAAGCGGTCCAGCCA 207
78 pThrAlaSerLeuAsnLysGlnHis 86
  :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
206 TACAGTGTACAGGTAACCGTTACCAT 182
```







OM of: US-09-528-682-4 to: Pending\_Patents\_NA\_Main: \* out\_format : pfs

Date: Jun 18, 2002 9:44 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool/US09528682/runat\_18062002\_082444\_7874/app\_query.fasta\_1.689  
-DB=pending\_Patents\_NA\_Main -QPMT=fastap -SUFFIX=p2n.rnpgm  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -GAPEXT=4.500 -CGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -XGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALLGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pfs -NORW=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09528682.@CGN1\_1.8511 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT\_THREADS=1

Search information block:

Query: US-09-528-682-4

Query length: 110

Database: Pending\_Patents\_NA\_Main: \*

Database sequences: 21979536

Database length: -2067452561

Search time (sec): 7363.740000

score\_list:

Sequence	Strd Orig	zScore	EScore Len	Documentation
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-404-520-11975 +	83.50	168.19	6.44	33
/cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-340-236-152 +	81.00	168.88	5.89	159
/cgn2_6/ptodata/2/pna/US0612_COMB.seq:US-60-128-476-1482 +	81.00	168.88	5.89	15
/cgn2_6/ptodata/2/pna/US0608_COMB.seq:US-60-068-139-236 +	81.00	165.02	9.68	236
/cgn2_6/ptodata/2/pna/US0908_COMB.seq:US-60-082-116-168 +	81.00	165.01	9.68	237
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-596-002-31 +	81.00	132.47	628.83	6579
/cgn2_6/ptodata/2/pna/US0614_COMB.seq:US-60-140-121-31 +	81.00	132.47	628.83	6579
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-268958 +	79.50	180.99	1.25	
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-385-678-268958 +	79.50	180.99	1.25	
/cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-540-236-823 +	78.00	167.77	6.79	858
/cgn2_6/ptodata/2/pna/US0612_COMB.seq:US-60-128-476-713 +	78.00	167.77	6.79	858
/cgn2_6/ptodata/2/pna/US0606_COMB.seq:US-60-068-139-71 -	78.00	165.18	9.47	1118
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-596-002-15 +	78.00	133.43	555.89	2862
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-533-559-4769 +	77.00	159.76	18.99	15
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:PCT-US02-03987-9496 -	77.00	154.45	37.51	26
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-072-851-9496 -	77.00	154.45	37.51	26
/cgn2_6/ptodata/2/pna/US0906_COMB.seq:US-09-608-175-457 +	77.00	148.68	78.59	472
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-23849 -	77.00	135.64	418.87	1
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-8854 -	77.00	135.64	418.87	1
/cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:PCT-US97-06635-1 -	76.00	152.06	51.00	2619
/cgn2_6/ptodata/2/pna/US090_COMB.seq:PCT-US97-06635A-1 -	76.00	152.06	51.00	2619
/cgn2_6/ptodata/2/pna/US0636_COMB.seq:US-60-360-207-18907 +	76.00	143.29	157.03	6
/cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-847-513A-1 +	76.00	115.90	5.3e+03	1051
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-134-000-1761 +	75.50	153.20	44.02	206
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-614-221A-288 +	74.50	147.99	85.89	27
/cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031-312 +	74.50	94.23	8.5e+04	6664
/cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031B-312 +	74.50	94.23	8.5e+04	6664
/cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031C-312 +	74.50	94.23	8.5e+04	6664
/cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-335-032-12214 +	74.50	94.23	8.5e+04	6664
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-3251 +	74.00	163.92	11.14	47
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-19447 +	74.00	160.41	17.47	6
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-264-73223 +	74.00	160.41	17.47	6
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-739-449-4641 +	74.00	156.02	30.67	10
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-110-4641 +	74.00	156.02	30.67	10
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-59882 +	74.00	138.96	273.63	18

/cgn2_6/ptodata/2/pna/US0616_COMB.seq:US-60-168-139-2269 -	74.00	135.59	421.49	
/cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-514-000-377 -	74.00	120.46	2.9e+03	
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-22857 +	74.00	118.27	3.9e+03	
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-7050 +	74.00	118.27	3.9e+03	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-264-68501 +	74.00	100.82	3.6e+04	
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-739-449-213 +	74.00	100.82	3.6e+04	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-110-213 +	74.00	100.82	3.6e+04	
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-085-959-35 +	73.50	139.30	261.72	52
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-533-559-5535 +	73.00	159.04	20.82	
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-7179 +	73.00	158.67	21.85	
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-167-568-44 +	73.00	142.84	166.30	32
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-167-568A-44 +	73.00	142.84	166.30	32
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-206-942-48 +	73.00	142.84	166.30	32
/cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-167-568-42 +	73.00	142.79	167.46	32
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-206-942-46 +	73.00	142.79	167.46	32
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-287-618-3163 +	72.50	160.92	16.37	
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-436-914A-2244 +	72.50	160.92	16.37	
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-561-875A-2244 +	72.50	160.92	16.37	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-811-107-71 +	72.50	153.02	45.07	4
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-565-309A-18506 +	72.00	156.68	28.20	
/cgn2_6/ptodata/2/pna/US085_COMB.seq:US-08-545-328-1 -	72.00	89.60	1.5e+05	580
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-021-323-13873 +	71.50	155.77	31.67	
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-54527 -	71.50	129.29	944.85	
/cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-1733 -	71.50	129.29	944.85	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-264-69026 -	71.00	162.85	12.78	
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-140827 +	71.00	162.85	12.78	
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-140827 +	71.00	162.85	12.78	
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-385-678-140827 +	71.00	137.81	316.82	3300
/cgn2_6/ptodata/2/pna/US07_COMB.seq:US-07-895-367-6 +	71.00	137.81	316.82	330
/cgn2_6/ptodata/2/pna/US081_COMB.seq:US-08-134-397A-6 +	71.00	137.81	316.82	330
/cgn2_6/ptodata/2/pna/US086_COMB.seq:US-08-614-377-6 +	71.00	137.81	316.82	330
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-379-931-6 +	71.00	137.81	316.82	330
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-565-306-61602 +	70.50	152.04	51.10	
/cgn2_6/ptodata/2/pna/US0631_COMB.seq:US-60-312-544-3131 +	70.50	147.22	94.78	
/cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:PCT-US02-03987-4742 +	70.50	137.38	335.18	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-242-4742 +	70.50	137.38	335.18	
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-072-851-4742 +	70.50	137.37	335.59	
/cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:PCT-US02-03987-8740 +	70.50	137.37	335.59	
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-242-8740 +	70.50	137.37	335.59	
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-072-851-8740 +	70.50	122.61	2.2e+03	1
/cgn2_6/ptodata/2/pna/US087_COMB.seq:US-08-781-986A-54 -	70.50	122.61	2.2e+03	13
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171-54 -	70.50	122.61	2.2e+03	1
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171B-54 -	70.50	122.61	2.2e+03	1
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171E-54 -	70.50	122.61	2.2e+03	1
/cgn2_6/ptodata/2/pna/US089_COMB.seq:US-08-956-171F-54 -	70.50	122.61	2.2e+03	1
/cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-654-617-135238 +	70.00	151.98	51.49	
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-135238 +	70.00	151.98	51.49	
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-4939 +	70.00	146.80	100.06	
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-305-638-18 -	70.00	137.64	323.79	26
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-305-638-12 -	70.00	135.68	416.56	32
/cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-525-160-4 -	70.00	135.68	416.56	32
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-305-638-11 -	70.00	133.67	474.48	40
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-620-392-59882 +	70.00	130.09	853.44	
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-521-017B-865 -	70.00	103.07	2.7e+04	







[illegible]







/cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:US-09-703-708-8292 + 65.00 124.12 1.8e+03 30  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-164-320-8292 + 65.00 124.12 1.8e+03 30  
/cgn2\_6/ptodata/2/pna/US0618\_COMB.seq:US-60-183-791-8292 + 65.00 124.12 1.8e+03 30  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:PCT-US97-21976-40 + 65.00 122.44 2.3e+03 3655  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-308-967-40 + 65.00 122.44 2.3e+03 3655  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-308-967-40 + 65.00 122.44 2.3e+03 3655  
/cgn2\_6/ptodata/2/pna/US0636\_COMB.seq:US-60-360-207-31471 + 65.00 117.21 4.5e+03 6  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-167-217-17696 - 65.00 113.66 7.0e+03 8  
/cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:US-09-902-540-1006 + 65.00 109.83 1.1e+04 13  
/cgn2\_6/ptodata/2/pna/US086\_COMB.seq:US-08-621-425-243 - 65.00 105.50 2.0e+04 2061  
/cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:US-09-703-708-710 + 65.00 101.22 3.5e+04 319  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-164-320-710 + 65.00 101.22 3.5e+04 319  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-183-791-710 + 65.00 101.22 3.5e+04 319  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-304-517A-79862 + 64.50 142.69 169.59 4  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-371-146A-79862 + 64.50 142.69 169.59 4  
/cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:US-09-565-306-152 + 64.50 142.69 169.59 409  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-654-617-290956 + 64.50 142.69 169.59  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-684-617-290956 + 64.50 142.69 169.59  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-985-678-79862 + 64.50 142.69 169.59  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-692-257A-6207 + 64.50 142.22 180.07 4  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-162-747-5733 + 64.50 142.22 180.07 4  
/cgn2\_6/ptodata/2/pna/US093\_COMB.seq:US-09-394-745-19552 + 64.50 142.11 182.71 4  
/cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:US-09-565-306-41349 + 64.50 142.11 182.71 4  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-60-220-535-3885 + 64.50 142.06 183.77 43  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-162-747-3109 + 64.50 142.04 184.30 43  
/cgn2\_6/ptodata/2/pna/US092\_COMB.seq:US-09-227-586-165 + 64.50 141.56 196.03 459  
/cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:US-09-976-054-165 + 64.50 141.56 196.03 459  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-654-617-406977 + 64.50 141.12 207.35  
/cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:US-09-669-817A-20788 + 64.50 141.12 207.35  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-684-016-40678 + 64.50 141.12 207.35  
/cgn2\_6/ptodata/2/pna/US0919\_COMB.seq:US-60-197-872-42677 + 64.50 140.92 212.79 4  
/cgn2\_6/ptodata/2/pna/US0617\_COMB.seq:US-60-172-363-1122 - 64.50 140.06 237.62 53  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-09-971-146A-98808 + 64.50 139.56 253.34 5  
/cgn2\_6/ptodata/2/pna/US0620\_COMB.seq:US-60-208-063-18026 + 64.50 139.44 257.30  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-09-865-413A-4093 + 64.50 139.44 257.30  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-60-208-063-24135 + 64.50 139.12 268.11  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-09-865-439A-68525 + 64.50 138.93 274.42  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-60-207-458-112685 + 64.50 138.93 274.42  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-09-865-439A-92792 + 64.50 138.74 281.33  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-60-207-458-136807 + 64.50 138.74 281.33  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-09-865-439A-64398 + 64.50 138.68 283.64  
/cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:US-60-207-458-108558 + 64.50 138.68 283.64  
/cgn2\_6/ptodata/2/pna/US0620\_COMB.seq:US-09-865-439A-98243 + 64.50 138.36 295.25  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-60-207-458-142434 + 64.50 138.36 295.25  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-680-598A-1072 - 64.50 137.10 347.43 7  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-974-300-1965 - 64.50 135.50 446.24 85  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-644-868-10062 - 64.50 135.25 440.11 8  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-654-617-387248 - 64.50 132.20 651.19  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-684-016-387248 - 64.50 132.20 651.19  
/cgn2\_6/ptodata/2/pna/US091\_COMB.seq:US-09-152-588-4 + 64.50 130.91 767.55 1361  
/cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:US-09-654-617-259159 + 64.50 127.49 1.2e+03 1931  
/cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:US-09-684-016-236363 + 64.50 127.24 1.2e+03 1931  
/cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:US-09-611-529-3363 + 64.50 126.81 1.3e+03 20  
/cgn2\_6/ptodata/2/pna/US0617\_COMB.seq:US-60-173-386-1797 + 64.50 126.71 1.3e+03 20  
/cgn2\_6/ptodata/2/pna/US0617\_COMB.seq:US-60-175-871-2007 + 64.50 126.71 1.3e+03 20  
/cgn2\_6/ptodata/2/pna/US0618\_COMB.seq:US-60-184-775-1834 + 64.50 126.71 1.3e+03 20  
/cgn2\_6/ptodata/2/pna/US0616\_COMB.seq:US-60-167-324-2000 + 64.50 126.70 1.3e+03 20  
/cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:US-60-164-150-42671 + 64.50 126.47 1.4e+03 2  
/cgn2\_6/ptodata/2/pna/US0619\_COMB.seq:US-60-191-637-42311 + 64.50 126.47 1.4e+03 2  
/cgn2\_6/ptodata/2/pna/US091\_COMB.seq:US-09-191-700-1955 + 64.50 126.47 1.4e+03 21  
/cgn2\_6/ptodata/2/pna/US091\_COMB.seq:US-09-107-532-3345 + 64.50 126.23 1.4e+03 219  
/cgn2\_6/ptodata/2/pna/US092\_COMB.seq:US-09-252-691C-2251 + 64.50 125.93 1.4e+03 221  
/cgn2\_6/ptodata/2/pna/US092\_COMB.seq:US-09-252-691C-2251 + 64.50 125.93 1.5e+03 221  
/cgn2\_6/ptodata/2/pna/US0925\_COMB.seq:US-60-250-830-2363 + 64.50 125.86 1.5e+03 22  
/cgn2\_6/ptodata/2/pna/US0632\_COMB.seq:US-60-323-966-2363 + 64.50 125.86 1.5e+03 22  
/cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:US-09-614-221A-1896 + 64.50 125.23 1.6e+03 24  
/cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:US-09-917-800A-1696 + 64.50 124.15 1.8e+03 2  
/cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:US-09-614-150-731 + 64.50 123.75 1.9e+03 283  
/cgn2\_6/ptodata/2/pna/US0617\_COMB.seq:US-60-173-464-623 + 64.50 123.75 1.9e+03 283



[illegible]



/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-306-349-25075 +	63.50	144.50	134.34	26
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-960-481-25075 +	63.50	144.50	134.34	27
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517-117862 +	63.50	143.97	143.92	28
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-117862 +	63.50	143.97	143.92	29
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-117862 +	63.50	143.97	143.92	30
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-241065 +	63.50	142.36	176.79	31
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-241065 +	63.50	142.36	176.79	32
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-97 +	63.50	142.36	176.79	33
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-241065 +	63.50	142.36	176.79	34
/cgn2_6/ptodata/2/pna/US06014_COMB.seq:US-60-144-08A-9961 +	63.50	142.36	176.79	35
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-239799 +	63.50	141.54	196.45	36
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-239799 +	63.50	141.54	196.45	37
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-2057 +	63.50	141.54	196.45	38
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-239799 +	63.50	141.54	196.45	39
/cgn2_6/ptodata/2/pna/US06014_COMB.seq:US-60-144-08A-10841 +	63.50	141.54	196.45	40
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-565-240-49764 +	63.50	141.06	208.86	41
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-240997 +	63.50	140.76	217.22	42
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-240997 +	63.50	140.76	217.22	43
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-222 +	63.50	140.76	217.22	44
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-240997 +	63.50	140.76	217.22	45
/cgn2_6/ptodata/2/pna/US06014_COMB.seq:US-60-144-08A-10162 +	63.50	140.76	217.22	46
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-362-510-34454 +	63.50	140.48	224.94	47
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-362-510A-34454 +	63.50	140.48	224.94	48
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-904-013-34454 +	63.50	140.48	224.94	49
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-245309 +	63.50	140.20	233.43	50
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-5514 +	63.50	140.20	233.43	51
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-245309 +	63.50	140.20	233.43	52
/cgn2_6/ptodata/2/pna/US06014_COMB.seq:US-60-144-08A-37112 +	63.50	140.20	233.43	53
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-5355 +	63.50	140.17	234.14	54
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-5355 +	63.50	140.17	234.14	55
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-304-517A-276676 +	63.50	140.05	237.70	56
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-276676 +	63.50	140.05	237.70	57
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-440-687-30183 +	63.50	140.05	237.70	58
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-276676 +	63.50	140.05	237.70	59
/cgn2_6/ptodata/2/pna/US06014_COMB.seq:US-60-144-08A-37010 +	63.50	140.05	237.70	60
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-874-708A-40123 -	63.50	139.42	257.81	61
/cgn2_6/ptodata/2/pna/US06021_COMB.seq:US-60-211-7-30679 -	63.50	139.42	257.81	62
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-565-309A-14598 +	63.50	138.68	283.41	63
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-146A-245309 +	63.50	138.17	302.74	64
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-634-306B-268304 +	63.50	137.88	314.01	65
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-439A-102126 -	63.50	137.25	340.62	66
/cgn2_6/ptodata/2/pna/US06020_COMB.seq:US-60-207-458-146064 -	63.50	137.25	340.62	67
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-565-309A-14599 +	63.50	136.92	355.20	68
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-565-309A-48145 +	63.50	136.90	355.02	69
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-5354 +	63.50	136.80	360.67	70
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-439A-25538 +	63.50	136.67	366.88	71
/cgn2_6/ptodata/2/pna/US06020_COMB.seq:US-60-207-458-69835 +	63.50	136.67	366.88	72
/cgn2_6/ptodata/2/pna/US06017_COMB.seq:US-60-172-371-9350 +	63.50	135.91	404.62	73
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-5354 +	63.50	135.82	409.39	74
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-5354 +	63.50	135.82	409.39	75
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-439A-98305 +	63.50	135.77	411.78	76
/cgn2_6/ptodata/2/pna/US06020_COMB.seq:US-60-207-458-142226 +	63.50	135.77	411.78	77
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-19654 +	63.50	135.55	423.76	78
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-19654 +	63.50	135.55	423.76	79
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-252-691-1442 +	63.50	135.17	444.70	80
/cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-985-678-1442 +	63.50	135.17	444.70	81
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-627 -	63.50	135.14	446.32	82
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-627 -	63.50	135.14	446.32	83
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-634-308B-201896 -	63.50	134.79	466.67	84
/cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-634-308B-201697 -	63.50	134.79	466.67	85
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-060-063-6149 -	63.50	131.84	682.08	86
/cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-533-559-4762 -	63.50	131.09	750.75	87
/cgn2_6/ptodata/2/pna/US06036_COMB.seq:US-60-360-207-44004 +	63.50	129.78	888.21	88
/cgn2_6/ptodata/2/pna/US06031_COMB.seq:US-60-316-362-2019 -	63.50	129.25	950.14	89
/cgn2_6/ptodata/2/pna/US06031_COMB.seq:US-60-316-362-1019 -	63.50	127.30	1.2e+03	90
/cgn2_6/ptodata/2/pna/US06016_COMB.seq:US-60-168-138-976 +	63.50	126.97	1.3e+03	91
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-263015 +	63.50	125.89	1.4e+03	92
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-263015 +	63.50	125.89	1.4e+03	93
/cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-505-532-4639 -	63.50	125.02	1.6e+03	94
/cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-819-019A-4639 -	63.50	125.02	1.6e+03	95
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-371-168-6555 -	63.50	124.53	1.7e+03	96
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-644-868-7963	63.50	124.53	1.7e+03	97



```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 2005-001
; CURRENT APPLICATION NUMBER: US/09/540, 236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 152
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-152

alignment_scores:
    Quality: 81.00      Length: 101
    Ratio: 1.397      Gaps: 6
Percent Similarity: 57.426      Percent Identity: 30.693

alignment_block:
US-09-528-682-4 x US-09-540-236-152 ..

Align seg 1/1 to: US-09-540-236-152 from: 1 to: 1596

10 AlaTyrGluProIleAlaAsnThrAsnThrThrThrGlnPheLeu...As 25
|||||:||||| :: ||||||| ||
973 GCCTATCATCCAACAAGCTGCACAAATACGCTGGGCGAATATTTGCAAAA 1022

25 nMetGlyAsn.....GluValAlaLeuAspGlyArgTyrSerAsnT 39
|||||:||||| :: ||||||| ||
1023 TTTGGGTAAAAACCCCACTTTGAATTTGCCGAAACCCGAAATACGCACATG 1072

39 yrAlaLeuIleSerAlaGluGly.....GlyMetAsp 49
|||||:||||| :: ||| |||
1073 TAACCTTTTCTTAGCGGTGCTGCTGAAGCGCTGCATCATGCGGAAGAT 1122

50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGl 66
|||||:||||| :: ||| |||:|
1123 CGCAATTTTGATTAAATAGT.....CCTGATGTTAAAC 1154

66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83
|||||:||||| :: ||| |||:|
1155 ATACGACCTAAACCTGAATGAGTCTTTTGAGGTTACCGAAAAATTA 1204

83 snLysGlnHisAspAla...AspPheLysTyrIleLysLeuLeuIle 98
|| ||||| ||| ||||| |||:|
1205 AT.....GATCGGATTGACTCGGCGCAATATGATGATTATTAATGTC 1245

99 Asn 99
|||
1246 AAT 1248

seq_name: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-60-128-476-1482

seq_documentation_block:
; Sequence 1482, Application US/60128476
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA
; TITLE OF INVENTION: CATARRHALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,139
; CURRENT APPLICATION NUMBER: US/60/128,476
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 5002
; SEQ ID NO 1482
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-60-128-476-1482

```



```

; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2369 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: MCALC236
US-60-068-139-236

alignment_scores:
    Quality: 81.00      Length: 101
    Ratio: 1.397      Gaps: 6
    Percent Similarity: 57.426      Percent Identity: 30.693

alignment_block:
US-09-528-682-4 x US-60-068-139-236

Align seg 1/1 to: US-60-068-139-236 from: 1 to: 2369

10 AlaTyrGluProIleAlaAsnThrAsnThrThrGlnPheLeu...As 25
|||||:||||: ||: |||||||: ||: |||||: ||: ||: ||: ||: ||
979 GCCTATCATCCAACAGTCTGACAATAGCTGGCGGAATATTTACAAA 1028
25 nMetGlyAsn.....GluValAlaLeuAspGlyArgTyrSerAsnT 39
1029 TTTGGGTAAACCCCAACTTAGAATTGCGGAACCGAAATAACGCACATG 1078
39 yrAlaLeuIleSerAlaGluGly.....GlyMetAsp 49
|||||:||||: ||: |||||: ||: |||||: ||: ||: ||: ||: ||
1079 TAACCTTTTCTTTAGTGGTGGTCTGAAGATCTGTACAACGGCGAAGAT 1128
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG 66
|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
1129 CGCATTTTGATTAAATAGT.....CCTGATGTTAAAAAC 1160
66 upheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83
|||||:||||: |||||: |||||: |||||: |||||: |||||: |||||:
1161 ATACGACCTAAAGCTGAATGAGTGCTTTTGAGGTACCGAAAAATTTAA 1210
83 snLysGlnHisAspAla...AspPheLysLysTyrIleLysLeuLeuIle 98
|| |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
1211 AT.....GATCGGATTGACTCGGCGCAATATCATCTATTAAATGTC 1251

99 Asn 99
|||
1252 AAT 1254

seq_name: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.us-60-082-116-166

seq_documentation_block:
; Sequence 168, Application US/60082116
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HANN, AMY L.
; APPLICANT: HEATH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; APPLICANT: BROOKS, JACQUELINE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATENA
; NUMBER OF INVENTION: 385
; NUMBER OF SEQUENCES: 385
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

```















```
;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068.139
; FILING DATE: HEREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0008-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: MCA1C071
; US-60-068-139-71

alignment_scores:
  Quality: 78.00 Length: 116
  Ratio: 1.300 Gaps: 6
  Percent Similarity: 51.724 Percent Identity: 29.310

alignment_block:
  US-09-528-682-4 x US-60-068-139-71/rev ..
  Align seg 1/1 to reverse of: US-60-068-139-71 from: 1 to: 1118

5 ValArgIleGlnAspAlaTyrGluProIleAla.....AsnThrAs 18
||||| :::::::::: :::::::::: :::::::::: ::::::::::
705 GTTACGACTCGTGATGCACCTAAATATGATTGCTAAATTCGTCACCAACCAA 656

18 nThrThrThrGlnPheLeuAsnMetGly.....AsnGluValAlaLeuA 33
||||| ||| |||||| ||| ||| ::::
655 CCAACACAGCCCTGTTTAGTCATGGTTATTAAACCTGTTGAATCA 606

33 sPgLYArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyMetAsp 49
||||| :::::::::: :::::::::: :::::::::: ::::::::::
605 TTGGT...TATGATAATTTTGTAAATTTGCTGCTGATGCTGCGGTTGAT 559

50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG1 66
||||: :::::
558 GCAATTTTGATG.....GTGGAT.....541

66 uPheAsnSerLeuProAsnAlaSerSerAspThrAlaSerLeuA 83
||||| ||:::||||: ::::: :::::
540 .....TTACCGCCCAAGAGCAGATGGCTATGCAAAATCACTGA 501

83 snLYsGlnHisAspAlaAspPheLysTyrIle.....94
::::: ||| ::::: :::::
500 CCGAGGTACCGATCATCCATGAATCAATTTTTCGTCGCCAACCAACC 451

95 .....LysLeuLeuIleAsnAsnAspGlyPhe 103
::::: ||| ::::: ||| :::::
450 ACAAATCAGATCGCGGTACCAAGGTTATCGAAACACTGATGTTGTTT 403

seq_name: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-60-082-116-250
seq_documentation_block:
; Sequence 250, Application US/60082116
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HANN, AMY L.
; APPLICANT: HEATH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; APPLICANT: BROOKS, JACQUELINE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 385
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/082.116
; FILING DATE: HEREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0008-3 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 250:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8860 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: MCA1C250
; US-60-082-116-250

alignment_scores:
  Quality: 78.00 Length: 116
  Ratio: 1.300 Gaps: 6
  Percent Similarity: 51.724 Percent Identity: 29.310

alignment_block:
  US-09-528-682-4 x US-60-082-116-250 ..
  Align seg 1/1 to: US-60-082-116-250 from: 1 to: 8860

5 ValArgIleGlnAspAlaTyrGluProIleAla.....AsnThrAs 18
||||| :::::::::: :::::::::: :::::::::: ::::::::::
8032 GTTACGACTCGTGATGCACCTAAATATGATTGCTAAATTCGTCACCAACCAA 8081

18 nThrThrThrGlnPheLeuAsnMetGly.....AsnGluValAlaLeuA 33
||||| ||| |||||| ||| ||| ::::
8082 CCAACACAGCCCTGTTTAGTCATGGTTATTAAACCTGTTGAATCA 8131

33 sPgLYArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyMetAsp 49
||||| :::::::::: :::::::::: :::::::::: ::::::::::
8132 TTGGT...TATGATAATTTTGTAAATTTGCTGCTGATGCTGCGGTTGAT 8178

50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG1 66
```







```

; Sequence 4769, Application US/095333559
;
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,621
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FASTSEQ for Windows Version
; SEQ ID NO 4769
; LENGTH: 1523
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
; US-09-533-559-4769

```

```
alignment_scores:
  Quality: 77.00      Length: 119
  Ratio: 1.167       Gaps: 4
  Percent Similarity: 55.462  Percent Identity: 23.529
```

alignment\_block:

US-09-528-682-4 x US-09-533-559-4769

Align seg 1/1 to: US-09-533-559-4769 from: 1 to: 1523

[illegible]

seq\_name: /cqn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:PCT-US02-03987-9496

```
seq_documentation_block:
; Sequence 9496, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying
```

```

; TITLE OF INVENTION: Proliferon
; FILE REFERENCE: ELITRA_028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9496
; LENGTH: 2619
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2619)
PCT-US02-03987-9496

```

```
alignment_scores:      Length: 94
                       Quality: 77.00
                       Ratio: 1.481
                       Gaps: 5
                       Percent Similarity: 55.319
                       Percent Identity: 29.787
```

alignment\_block:

US-09-528-682-4 x PCT-US02-03987-9496/rev

Align seq 1/1 to reverse of: PCT-US02-03987-9496 from: 1 to: 2619

```

14  lLeAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
    :::::::::: ::::::::::::::::::::
2028 GTTGCCAAACATGAGTACCACCAAAAGTTC.....AATAGAGT 1991

30 lAlaLeuAspGlyArgTyrSerAsnTyr...AlaLeuIleSerAlaGluG 46
    l::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
1990 AGTCACCGATAGTCACGACACGAACATTCCTTCGCGTATTCTCACCAAAAG 1941

46 lYcGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
    llll ::::: ::::: ::::: ::::: ::::: :::::
1940 AGGGCGCATAGTCGCCATTCTTTTAGCAGTGTCAATATC..... 1903

63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
    ::::::::::CGTTTCAAC..... 1894

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysTyrIleLysL 96
    llllll :::: llllll ::::: ::::: :::::
1893 ....TGCTTTAACTCAAGAGCTTCCCAGATT.....TTCGTTTGACT 1854

96 euLeuIleAsnAsnAspGlyPhePheSerAsn 106
    ::::::::::::::: ::::::::::::::: :::::::::::::::
1853 TGCTGTGTTCAATCGCAGCGCAATTTCTTCACAGAGT 1822

```

seq\_name: /cqn2\_6/ptodata/2/pna/US098A\_COMB.seq:US-09-815-242-9496

```
seq_documentation_block:
  : Sequence 9496, Application US/09815242
  :
  : GENERAL INFORMATION:
  :
  : APPLICANT: Haselbeck, Robert
  : APPLICANT: Ohlssen, Kari L.
  : APPLICANT: Zyskind, Judith W.
  : APPLICANT: Wall, Daniel
  : APPLICANT: Trawick, John D.
  : APPLICANT: Carr, Grant J.
  : APPLICANT: Yamamoto, Robert T.
  : APPLICANT: Xu, H. Howard
  :
  : TITLE OF INVENTION: Identification of
  :
  : FILE OF INVENTION: Prokaryotes
  :
  : FILE REFERENCE: ELTRA 011A
  :
  : CURRENT APPLICATION NUMBER: US/09/815
  :
  : CURRENT FILING DATE: 2003-03-21
  :
  : PRIOR APPLICATION NUMBER: 60/191,078
  :
  : PRIOR FILING DATE: 2000-03-21
  :
  : PRIOR APPLICATION NUMBER: 60/206,848
  :
```



:	APPLICANT:	Cart, Grant J.
:	APPLICANT:	Xu, H. Howard
:	APPLICANT:	Foulkes, J. Gordon
:	APPLICANT:	zamudio, Carlos
:	APPLICANT:	Haselbeck, Robert
:	APPLICANT:	Ohlsen, Karl L.
:	APPLICANT:	Zyskind, Judith W.
:	APPLICANT:	Wall, Daniel
:	APPLICANT:	Trawick, John D.
:	APPLICANT:	Yanamoto, Robert T.
:	APPLICANT:	Roemer, Terry

TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE  
 TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF  
 NUMBER OF SEQUENCES: 1175  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE



```
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,175
; FILING DATE: HEREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0009-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-416
; INFORMATION FOR SEQ ID NO: 457:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4720 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: SPN2c459
; US-60-068-175-457

alignment_scores:
  Quality: 77.00 Length: 94
  Ratio: 1.481 Gaps: 5
  Percent Similarity: 55.319 Percent Identity: 29.787

alignment_block:
US-09-528-682-4 x US-60-068-175-457 ..
Align seg 1/1 to: US-60-068-175-457 from: 1 to: 4720
14 lIeAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
   ::::: ::::: ::::: ::::: :::::
2169 GTTGCCAAACATGATGACGACCAACAAAGTTC.....AATAGACT 2206
30 lAlaLeuaspGlyArgTyrSerAsnTyr...AlaLeuIleSerAlaGluG 46
   ::::: ::::: ::::: ::::: :::::
2207 AGTCACCGATAGTCACGACACGAACTTCCTGCCGTATTTTCACCAAG 2256
46 lYgLYmetAspArgAspLeuPheGlySerAlaAsnIleaspGlyPhePro 62
   ::::: ::::: ::::: ::::: :::::
2257 AGGGCATAGCTCCCATCTCTTAGCAGTGCAATATC..... 2294
63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
   ::::: :::::
2295 .....CGTTTCAAC..... 2303
79 rAlaSerLeuAsnLysGlnHisAspAlaaspPheLysTyrIleLysL 96
   ::::: :::::
2304 ....TGCTTAACCTCAAGAGCTTCCAGATTTT.....TTCGTTGACT 2343
96 eUleuIleAsnAsnAspGlyPhePheSerAsn 106
   ::::: :::::
2344 TGCTGTTCAATCGACGCAATCTTCAGCAGT 2375

seq_name: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-23849

seq_documentation_block:
; Sequence 23849, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 23849
; LENGTH: 17891
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-620-392-23849

alignment_scores:
  Quality: 77.00 Length: 111
  Ratio: 1.481 Gaps: 5
  Percent Similarity: 46.847 Percent Identity: 28.829

alignment_block:
US-09-528-682-4 x US-09-620-392-23849/rev ..
Align seg 1/1 to reverse of: US-09-620-392-23849 from: 1 to: 17891
46 GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleaspGlyPhePr 62
   ::::: ::::: ::::: ::::: :::::
5124 GGTTCCTACCTAGGAATCTTGCTGGTGACAAGGTACACCGTGCCTCC 5075
62 oGluValArgGluPhe...AsnSerLeuProAsnAsnLysAlaSerSera 78
   ::::: ::::: ::::: ::::: :::::
5074 AGAGGTTGTAGAGTTCTATCAAAAGTCTCATGAAGCGTGAAGCAAGAAG 5025
78 spThraLeuSerLeu.AsnLysGlnHis..... 86
   ::::: ::::: ::::: ::::: :::::
5024 ACACAACTTCTCTGGGATCAACAACATCAAGTCTCTCTGATGAGAAGC 4975
86 ..... 86
4974 AACATGATTGGAGAGATTGAGAAATAGATCAACATTCCTCTTAGCTGAAG 4925
87 .....AspAlaaspPheLysLysTyrIle..... 94
   ::::: ::::: ::::: :::::
4924 TTATTCAAGATGATGCTACCTTTTCAGAAATATATACTACTGAATTCCTTT 4875
95 ...LysLeuLeuIleAsnAsnasp..... 101
   ::::: ::::: ::::: :::::
4874 AGTCCACTGCTAATTCAGGTAGATATTATTATGAGGACTTACTTTTATATA 4825
102 .....GlyPhePheSerAsnAsnGlyGly 109
   ::::: ::::: ::::: :::::
4824 AAAAAATGGATTGTTCAGGTCAAAAGTGA 4794

seq_name: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-702-134-8854

seq_documentation_block:
; Sequence 8854, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 8854
; LENGTH: 17891
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-702-134-8854
```



```
alignment_scores:
  Quality: 77.00      Length: 111
  Ratio: 1.481       Gaps: 5
  Percent Similarity: 46.847  Percent Identity: 28.829

alignment_block:
US-09-528-682-4 x US-09-702-134-8854/rev ..
Align seg 1/1 to reverse of: US-09-702-134-8854 from: 1 to: 17891

46 GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePr 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5124 GGTTCCTACTAGGATCTTCTGCTGGTGGACAGGTACACCGTCTCC 5075

62 oGluValAlaArgGluPhe...AsnSerLeuProAsnAsnLysAlaSerSera 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5074 AGAGGTTGTAGAGTTCTATCAAAAGTCTCATGAAGCGTGAAGCCAAGAGG 5025

78 spThrAlaSerLeu.AsnLysGlnHis..... 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5024 ACACAACCTTCTGGGATCAACAACATCAAGTGTCTCTGATGTGAGAAGC 4975

86 ..... 86
4974 AACATGATTGGAGATTGAGATAGATCAACATTCCTCTTAGCTGTAAAG 4925

87 .....AspAlaAspPheLysLysTyrIle..... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4924 TTATTCAGATGATGCTACCTTTTCAGAAATATATATACTACTGAATTCCTTT 4875

95 ...LysLeuLeuIleAsnAsnAsp..... 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4874 AGTCCACTGCTAATTCAGGTAGATTATATGAGGACTTACTTTTATATA 4825

102 .....GlyPhePheSerAsnAsnGlyGly 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4824 AAAAAATGGATTGTTTCAGGTCAAAGTGA 4794

seq_name: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-815-264-81462

seq_documentation_block:
; Sequence 81462, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 81462
; LENGTH: 17891
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81462

alignment_scores:
  Quality: 77.00      Length: 111
  Ratio: 1.481       Gaps: 5
  Percent Similarity: 46.847  Percent Identity: 28.829
```

```
alignment_block:
US-09-528-682-4 x US-09-815-264-81462/rev ..
Align seg 1/1 to reverse of: US-09-815-264-81462 from: 1 to: 17891

46 GlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePr 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5124 GGTTCCTACTAGGATCTTCTGCTGGTGGACAGGTACACCGTCTCC 5075

62 oGluValAlaArgGluPhe...AsnSerLeuProAsnAsnLysAlaSerSera 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5074 AGAGGTTGTAGAGTTCTATCAAAAGTCTCATGAAGCGTGAAGCCAAGAGG 5025

78 spThrAlaSerLeu.AsnLysGlnHis..... 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5024 ACACAACCTTCTGGGATCAACAACATCAAGTGTCTCTGATGTGAGAAGC 4975

86 ..... 86
4974 AACATGATTGGAGATTGAGATAGATCAACATTCCTCTTAGCTGTAAAG 4925

87 .....AspAlaAspPheLysLysTyrIle..... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4924 TTATTCAGATGATGCTACCTTTTCAGAAATATATATACTACTGAATTCCTTT 4875

95 ...LysLeuLeuIleAsnAsnAsp..... 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4874 AGTCCACTGCTAATTCAGGTAGATTATATGAGGACTTACTTTTATATA 4825

102 .....GlyPhePheSerAsnAsnGlyGly 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4824 AAAAAATGGATTGTTTCAGGTCAAAGTGA 4794

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US97-06635-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9706635
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06635
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607991.8
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31458-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
```



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
PCT-US97-06635-1

alignment_scores:
  Quality: 76.00 Length: 94
  Ratio: 1.490 Gaps: 5
  Percent Similarity: 54.255 Percent Identity: 30.851

alignment_block:
  US-09-528-682-4 x PCT-US97-06635-1/rev ..
  Align seg 1/1 to reverse of: PCT-US97-06635-1 from: 1 to: 2619
    14 IleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
    2028 GTTCCACACATGAGTACACACCAAGTTC.....AATAGAGT 1991
    30 lAlaLeuAspGlyArgTyrSerAsn...TyrAlaLeuIleSerAlaGluG 46
    1990 AGTCACCGATAGTCACGACACGAAGTTCGCGTATTCTCACCAGAAG 1941
    46 lyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
    1940 AGGCGCATAGTCCCATTTCTTTAGCAGTGTCAATATC.....1903
    63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
    1902 .....CGTTTCAAC.....1894
    79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysTyrIleLysL 96
    1893 ....TGCTTTAACTTCAAGAGCTTCCAGATTTT.....TTCGTTGACT 1854
    96 euLeuIleAsnAsnAspGlyPhePheSerAsn 106
    1853 TGCTGTTCATCGCAGCGCAATTTCTTCAGCAGT 1822
  seq_name: /cgn2_6/ptodata/2/pna/PTUS_COMB.seq:PCT-US97-06635A-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9706635A
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06635A
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607991.8
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31458-5
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
PCT-US97-06635A-1

alignment_scores:
  Quality: 76.00 Length: 94
  Ratio: 1.490 Gaps: 5
  Percent Similarity: 54.255 Percent Identity: 30.851

alignment_block:
  US-09-528-682-4 x PCT-US97-06635A-1/rev ..
  Align seg 1/1 to reverse of: PCT-US97-06635A-1 from: 1 to: 2619
    14 IleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluVa 30
    2028 GTTCCACACATGAGTACACACCAAGTTC.....AATAGAGT 1991
    30 lAlaLeuAspGlyArgTyrSerAsn...TyrAlaLeuIleSerAlaGluG 46
    1990 AGTCACCGATAGTCACGACACGAAGTTCGCGTATTCTCACCAGAAG 1941
    46 lyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
    1940 AGGCGCATAGTCCCATTTCTTTAGCAGTGTCAATATC.....1903
    63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
    1902 .....CGTTTCAAC.....1894
    79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysTyrIleLysL 96
    1893 ....TGCTTTAACTTCAAGAGCTTCCAGATTTT.....TTCGTTGACT 1854
    96 euLeuIleAsnAsnAspGlyPhePheSerAsn 106
    1853 TGCTGTTCATCGCAGCGCAATTTCTTCAGCAGT 1822
  seq_name: /cgn2_6/ptodata/2/pna/PTUS_COMB.seq:PCT-US97-06635A-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9706635A
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06635A
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607991.8
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31458-5
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
PCT-US97-06635A-1

```











```

96 euLeuileAsnAsnAspGlyPheSerAsnAsn 107
:||||| |||||: : : : : : |||
2498 TACTTAAATTAATGATACGTCATACAAAC 2532
seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031-312

seq_documentation_block:
; Sequence 312, Application US/09012031A
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 1107.73251
; CURRENT APPLICATION NUMBER: US/09/012,031A
; CURRENT FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: 60/035,917
; EARLIER FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 317
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 666448
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-012-031-312

alignment_scores:
Quality: 74.50 Length: 95
Ratio: 1.330 Gaps: 2
Percent Similarity: 58.947 Percent Identity: 26.316

alignment_block:
US-09-528-682-4 x US-09-012-031B-312 ..
Align seg 1/1 to: US-09-012-031B-312 from: 1 to: 666448

13 ProfileAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnG1 29
||| |||||: ||| : : : : : |||
481149 CCGAGTGAGACGTTAATTCACACACACTCTCTATTTTGCAGTCTTTAAAC... 481196

29 uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuileSerAlaGluG 46
: : : : : |||||: ||| : : : : : |||
481197 .ATGAGTCTTGATGGGAGAAAAAATCGTCCAGTCTACACGCAACAT 481245

46 lyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAspGlyPhePro 62
: : : : : ||| : : : : : |||
481246 CAGTGTACCTAGTACATAAGACAGACAAT..... 481277

63 GluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSerAspTh 79
: : : : : |||||: ||| : : : : : |||||
481278 .....CAGCATTTCAATGACATAAACAGATGCTAGGCGAGTAGTGACGA 481321

79 rAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysTyrIleLysL 96
: : : : : |||||: ||| : : : : : |||
481322 AGATGCCCTTCCCAAGGCCAATCATTAATTTCAATAAGAACTACCAA 481371

96 euLeuileAsnAsnAspGlyPheSerAsnAsn 107
:||||| |||||: : : : : : |||
481372 TACTTAAATTAATGATACGTCATACAAAC 481406

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-012-031C-312

seq_documentation_block:
; Sequence 312, Application US/09012031C
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 1107.73251
; CURRENT APPLICATION NUMBER: US/09/012,031C
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/035,917
; PRIOR FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 666448
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-012-031C-312

alignment_scores:
Quality: 74.50 Length: 95

```







```

; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 73323
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-73323

alignment_scores:
  Quality: 74.00      Length: 79
  Ratio: 1.542      Gaps: 5
  Percent Similarity: 60.759      Percent Identity: 34.177

alignment_block:
US-09-528-682-4 x US-09-815-264-73323 ..

Align seg 1/1 to: US-09-815-264-73323 from: 1 to: 684

10 AlatyGluPProile...AlaAsnThrAsnThrThrThrGlnPheLeuAs 25
||||| |||||| :||||: :||: :||||:|||||
166 GCTTTGGACCTATAGCTCAGAGATCAGGATGAGTAGCAATTTCTCGC 215
||| :||||| :||| ||||| |||:
25 nMetGlyAsnGluValAlaLeu.....AspGlyArgTyrSerAsnTyrA 40
||| :||||| :||| ||||| |||:
216 CATGATATCTCCTGCTGCTCTCCTCAGCCATGGCACCCTACAGC...TACA 262
||| :||||| :||| ||||| |||:
40 laLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAla 56
|| :||||| :||| ||||| |||:
263 GTGCCACCACAGCTTCAGGCTGGGAAGACCAAGACTTCTTCAGGCATTGC 312
||| :||||| :||| ||||| |||:
57 AsnIle.....AspGlyPheProGluValArgGluPheAsnSe 69
||| :||||| :||| ||||| |||:
313 CCAGTGTCCCGTGCAGCGCGTGGCAGACAGATCCGG..... 351

69 rLeuProAsnAsnIysAlaSerSerAspThrAlaSer 81
||||| :||||:
352 .TTCCCTCAGCGCCTGCAATCCAGCACACATCATCA 387

seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-739-449-4641

seq_documentation_block:
; Sequence 4641, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 4641
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-4641

alignment_scores:
  Quality: 74.00      Length: 108
  Ratio: 1.138      Gaps: 5
  Percent Similarity: 60.185      Percent Identity: 25.000

alignment_block:
US-09-528-682-4 x US-09-739-449-4641 ..

Align seg 1/1 to: US-09-739-449-4641 from: 1 to: 1071

5 ValArgIleGln...AspAlaTyrGluProIleAlaAsnThrAsnThrTh 20
||||| :||||: ||| ||||| :|||||

```



```
493 GTCCGCTCGAGTCGATGAATATAAGCGCATCCCAACACCGATTCGCT 542
20 rThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyr 37
||||| :
543 GACGCGC.....CTCTCCAAACCGCGCGCTTTCGACGACCGCTGG 583
: :
37 erAsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPhe 53
: :
584 CC.....TCCGTTATGACAGCTCCATCGGCTGCAATATACACGCTC 627
: :
54 GlySerAlaAsnIleAspGlyPheProGluValArgGlu....PheAsnSe 69
: :
628 GTCTTCTCGATATCGATCACTTCAAGCGAATCAACGATACGTTCCGCCA 677
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
: :
678 TCCGCTTCTCGATATCGATCACTTCAAGCGAATCAACGATACGTTCCGCCA 677
: :
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
: :
678 TCCGCTTCTCGATATCGATCACTTCCGCCACCGTCCGCTC..... 717
: :
86 isAspAlaAspPheLysTyrIleLysLeuLeuIleAsnAsnAspGly 102
||||| :
718 TCCGCTTCTCGATATCGATCACTTCCGCCACCGTCCGCTC..... 717
: :
86 isAspAlaAspPheLysTyrIleLysLeuLeuIleAsnAsnAspGly 102
||||| :
718 TCCGCTTCTCGATATCGATCACTTCCGCCACCGTCCGCTC..... 717
: :
103 PhePheSerAsnAsnGlyGlyLys 110
||| :
745 TTCTGTTCCCGGACGCGCGCGAG 768
```

seq\_name: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:US-09-803-110-4641

```
seq_documentation_block:
; Sequence 4641, Application US/09803110
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)D
; CURRENT APPLICATION NUMBER: US/09/803,110
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: US 60/168,139
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 4641
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-803-110-4641
```

```
alignment_scores:
  Quality: 74.00      Length: 108
  Ratio: 1.138       Gaps: 5
  Percent Similarity: 60.185      Percent Identity: 25.000
```

```
alignment_block:
US-09-528-682-4 x US-09-803-110-4641
..
Align seg 1/1 to: US-09-803-110-4641 from: 1 to: 1071
```

```
5 ValArgIleGln...AspAlaTyrGluProIleAlaAsnThrAsnThr 20
||||| :
493 GTCCGCTCGAGTCGATGAATATAAGCGCATCCCAACACCGATTCGCT 542
20 rThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyr 37
||||| :
543 GACGCGC.....CTCTCCAAACCGCGCGCTTTCGACGACCGCTGG 583
: :
37 erAsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPhe 53
: :
584 CC.....TCCGTTATGACAGCTCCATCGGCTGCAATATACACGCTC 627
```

```
54 GlySerAlaAsnIleAspGlyPheProGluValArgGlu....PheAsnSe 69
: :
628 GTCTTCTCGATATCGATCACTTCAAGCGAATCAACGATACGTTCCGCCA 677
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
: :
678 TCCGCTTCTCGATATCGATCACTTCCGCCACCGTCCGCTC..... 717
: :
86 isAspAlaAspPheLysTyrIleLysLeuLeuIleAsnAsnAspGly 102
||||| :
718 TCCGCTTCTCGATATCGATCACTTCCGCCACCGTCCGCTC..... 717
: :
103 PhePheSerAsnAsnGlyGlyLys 110
||| :
745 TTCTGTTCCCGGACGCGCGCGAG 768
```

seq\_name: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:US-09-620-392-59882

```
seq_documentation_block:
; Sequence 59882, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 59882
; LENGTH: 6118
; TYPE: DNA
; ORGANISM: Oryza sativa
; OTHER INFORMATION: unsure at all n locations
US-09-620-392-59882
```

```
alignment_scores:
  Quality: 74.00      Length: 79
  Ratio: 1.542       Gaps: 5
  Percent Similarity: 60.759      Percent Identity: 34.177
```

```
alignment_block:
US-09-528-682-4 x US-09-620-392-59882
..
Align seg 1/1 to: US-09-620-392-59882 from: 1 to: 6118
```

```
10 AlaTyrGluProIle...AlaAsnThrAsnThrThrThrGlnPheLeuAs 25
||||| :
4966 GTTTTTCACCTATAGAGTCAGAGATCAGGATGAGTAGAGTAATAATTCCTCGC 5015
25 nMetGlyAsnGluValAlaLeu.....AspGlyArgTyrSerAsnTyrA 40
||| :
5016 CATGTATCTCTGCTGCTCTCTCTCCAGCCATGGCACCTACAGC...TACA 5062
40 laLeuIleSerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAla 56
: :
5063 GTCCACACACAGCTTCAGGCTGGGAAGACCAAGACTTCTTCAGGCATGTC 5112
57 AsnIle.....AspGlyPheProGluValArgGluPheAsnSe 69
: :
5113 CCAGTGTCCCGGTCGACGCGAGTGGGACGAGATCCGG..... 5151
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSer 81
||||| :
5152 .TTCCCTCACGGCTGCAATCCAGCAACACATCATCA 5187
```

seq\_name: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:US-60-168-139-2269

```
seq_documentation_block:
; Sequence 2269, Application US/60168139
; GENERAL INFORMATION:
```



```

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(15490)A
; CURRENT APPLICATION NUMBER: US/60/168,139
; CURRENT FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 3432
; SEQ ID NO 2269
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
US-60-168-139-2269

alignment_scores:
  Quality: 74.00      Length: 108
  Ratio: 1.138      Gaps: 5
  Percent Similarity: 60.185      Percent Identity: 25.000

alignment_block:
US-09-528-682-4 x US-60-168-139-2269/rev ..

Align seg 1/1 to reverse of: US-60-168-139-2269 from: 1 to: 8630

5 ValArgIleGln...AspAlaTyrgluProIleAlaAsnThrAsnThrTh 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5487 GTGGCCTCGAGTCGATGATATAGCGCATCGCCACACCGATTGCT 5438
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 rThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrS 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5437 GACGGCGC.....CTCTCCACCGCGCGCTTCGACGACCGCTGG 5397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
37 erAsnTyAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPhe 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5396 CC.....TCCGTTTATGACAGCTCCATCGCGCTGCAATATACCACGCTC 5353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 GlySerAlaAsnIleAspGlyPheProGluValArgGlu...PheAsnSe 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5352 GTCCTTCTGGATATCGATCATCTTCAAGCGAATCAACGATAGCTTCGGCCA 5303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5302 TCCGGTCGGCGACAGAATTCCTCGCCACCGTCGCTCCGTC..... 5263
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 isAspAlaAspPheLysLysTyrrileLysLeuLeuIleAsnAspGly 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5262 .....ATCCGCGCCAAATGTGCGCAAGGATGTT 5236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 PhePheSerAsnAsnGlyGlyLys 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5235 TTCGTTGCCCGAAGCGCGCGGAG 5212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

seq_name: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:us-09-514-000-377

seq_documentation_block:
; Sequence 377, Application US/09514000
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)B
; CURRENT APPLICATION NUMBER: US/09/514,000
; CURRENT FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 15034
; SEQ ID NO 377
; LENGTH: 40451
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
US-09-514-000-377

```

```

alignment_scores:
  Quality: 74.00      Length: 79
  Ratio: 1.138      Gaps: 5
  Percent Similarity: 60.185      Percent Identity: 25.000

alignment_block:
US-09-528-682-4 x US-09-514-000-377/rev ..

Align seg 1/1 to reverse of: US-09-514-000-377 from: 1 to: 40451

5 ValArgIleGln...AspAlaTyrgluProIleAlaAsnThrAsnThrTh 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30983 GTGGCCTCGAGTCGATGATATAGCGCATCGCCACACCGATTGCT 30934
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 rThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrS 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30933 GACGGCGC.....CTCTCCACCGCGCGCTTCGACGACCGCTGG 30893
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
37 erAsnTyAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPhe 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30892 CC.....TCCGTTTATGACAGCTCCATCGCGCTGCAATATACCACGCTC 30849
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 GlySerAlaAsnIleAspGlyPheProGluValArgGlu...PheAsnSe 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30848 GTCCTTCTGGATATCGATCATCTTCAAGCGAATCAACGATAGCTTCGGCCA 30799
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 rLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnH 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30798 TCCGGTCGGCGACAGAATTCCTCGCCACCGTCGCTCCGTC..... 30759
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 isAspAlaAspPheLysLysTyrrileLysLeuLeuIleAsnAsnAspGly 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30758 .....ATCCGCGCCAAATGTGCGCAAGGATGTT 30732
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 PhePheSerAsnAsnGlyGlyLys 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30731 TTCGTTGCCCGAAGCGCGCGGAG 30708
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

seq_name: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-22857

seq_documentation_block:
; Sequence 22857, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 22857
; LENGTH: 50625
; TYPE: DNA
; ORGANISM: Oryza sativa
; OTHER INFORMATION: unsure at all n locations
US-09-620-392-22857

alignment_scores:
  Quality: 74.00      Length: 79
  Ratio: 1.542      Gaps: 5
  Percent Similarity: 60.759      Percent Identity: 34.177

alignment_block:
US-09-528-682-4 x US-09-620-392-22857 ..

Align seg 1/1 to: US-09-620-392-22857 from: 1 to: 50625

10 AlaTyrgluProIle...AlaAsnThrAsnThrThrGlnPheLeuAs 25
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```















```
20 rThrGlnPheLeuAsnMet...GlyAsnGluVal..... 30
   :::::||||:::|||||
393 GCGGATTTGTTGAGGTGAGGCAATGATATTCGTGTTGTTGTAAC 442
   ||:::|||||
31 .....AlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIle 42
   ||:::|||||
443 GTGGAAGGAGAGGTGCAATGGAGGATCGTTATCT.....GCAC TTGTC 486
   ||:::|||||
43 SerAlaGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleAs 59
   :::::|||||
487 AATTTCAGAGGAGATTCAAAACAGGGTATTTTGGTGTATTTGATGGACA 536
   ||| |||:::|||||
59 pGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysAlas 76
   ||| |||:::|||||
537 TGGAGGACCTAAGCTGCAGAGTTTGCAGCAGACCACTTGAATGAAGAACA 586
   ||| |||:::|||||
76 erSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLys 92
   ||| |||:::|||||
587 TTATGGAT.....GAATTAGTAAGGAGGAATGATGAAGATGTTGTGGAA 630
   ||| |||:::|||||
93 TyrIleLys 95
   :::::|||||
631 GTGCTGAAA 639
```



OM of: US-09-528-682-4 to: Pending\_Patents\_NA\_New: \* out\_format : pfs

Date: Jun 18, 2002 9:49 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-Q=/cgn2\_1/uspto.spool/US09528682/runat\_18062002.082444\_7900/app\_query.fasta\_1.689  
-DB=Pending\_Patents\_NA\_New -QPMT=fastap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -FGAPOP=4.500 -FGAPEXT=0.050 -FGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682@cgn1\_1\_572 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-4

Query length: 110

Database: Pending\_Patents\_NA\_New: \*

Database sequences: 1014543

Database length: 727792371

Search time (sec): 403.930000

score\_list:

Sequence	Strid Orig	zScore	EScore	Len	Documentation
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-3042	79.50	177.47	0.0900	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-3340	78.00	166.42	0.3715	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-517	72.50	158.27	1.06	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-770-105A-3	69.50	161.69	0.6813	59
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-770-105A-1	69.50	157.40	1.18	91
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-041-018-103	69.50	145.93	5.14	2
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-721-544-8479	69.00	165.26	0.4309	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-919-002-2707	69.00	144.67	6.05	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	PCT-US02-09188-363	69.00	143.22	7.29	3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	PCT-US02-09370-379	69.00	143.22	7.29	3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-105-299-500	69.00	143.22	7.29	3
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-935-13145	69.00	139.18	12.23	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-25380	67.00	149.54	3.24	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-43631	67.00	144.35	6.30	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-43853	66.50	137.50	15.16	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-43498	66.00	152.18	2.31	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-43499	66.00	152.18	2.31	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-43500	66.00	152.18	2.31	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-43501	66.00	152.18	2.31	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-4674	66.00	150.12	3.00	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-19402	65.50	151.14	2.64	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-19403	65.50	151.14	2.64	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-17507	65.50	148.83	3.54	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-46339	65.50	137.04	16.09	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-3473	65.50	134.16	23.28	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-25817	65.50	126.99	58.39	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-42700	65.00	148.47	3.71	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-45623	64.50	134.28	22.93	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-539-3310-15353	64.00	155.37	1.53	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-3507	64.00	140.46	10.38	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-110883	64.00	133.69	24.71	2
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-110884	64.00	133.59	25.05	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-975-254-29497	63.50	154.22	1.78	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-268304	63.50	147.34	4.29	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-620-393B-4281	63.50	146.36	4.87	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-935-625-62553	63.50	146.13	5.01	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-201696	63.50	144.18	6.44	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-201697	63.50	144.18	6.44	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-2019	63.50	138.52	13.31	

/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-2019	63.50	138.52	13.31	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-1019	63.50	136.52	17.20	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-1019	63.50	136.52	17.20	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-7019	63.50	136.46	17.33	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-7019	63.50	136.46	17.33	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-6019	63.50	135.70	19.11	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-6019	63.50	135.70	19.11	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-19	63.50	128.19	50.02	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-19	63.50	128.19	50.02	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-5019	63.50	127.83	52.43	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-5019	63.50	127.83	52.43	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-437	63.00	147.59	4.16	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-26725	63.00	140.77	9.96	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-164418	63.00	140.71	10.05	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-09237-136	63.00	133.80	24.38	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-09370-210	63.00	133.80	24.38	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-105-299-1946	63.00	133.80	24.38	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-2498	63.00	128.74	46.66	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-2498	63.00	128.74	46.66	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-1498	63.00	128.04	51.02	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-1498	63.00	128.04	51.02	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-7498	63.00	127.56	54.23	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-7498	63.00	127.56	54.23	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-6498	63.00	126.74	60.30	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-6498	63.00	126.74	60.30	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-498	63.00	123.13	95.77	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-128-714-498	63.00	123.13	95.77	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-09370-1311	63.00	120.96	126.44	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-36917	62.50	145.55	5.40	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-36917	62.50	138.85	12.76	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-10-027-632-257992	62.50	136.48	17.28	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-540-209B-5017	62.50	129.91	40.15	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-45878	62.50	118.20	180.30	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-141-670-3	62.50	113.77	318.27	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-141-670-3	62.50	112.39	379.71	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-721-544-12765	62.00	147.83	4.03	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-33913	62.00	135.96	18.48	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-106-275-3	62.00	135.22	20.31	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-45008	62.00	135.17	20.44	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-107-907-4	62.00	129.65	41.48	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-107-828-4	62.00	129.65	41.48	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-45706	62.00	126.57	61.61	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-25215	61.50	139.42	11.86	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-25216	61.50	139.42	11.86	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-25801	61.50	136.67	16.88	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-46160	61.50	125.74	68.56	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-25318	61.50	125.50	70.63	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-114-544-1	61.50	125.24	73.07	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-25277	61.00	136.49	17.26	
/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq	US-60-360-039-34143	61.00	133.70	24.68	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-047-1387	61.00	127.35	55.71	
/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq	US-09-882-227-521	61.00	112.63	368.03	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-49885	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-49886	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-74683	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-74684	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-31388	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-31389	60.50	139.38	11.92	
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq	US-10-027-632-215384	60.50	138.00	14.21	
/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq	PCT-US02-13142-24550	60.50	136.15	18.03	



























```
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-379-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-383-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-383-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-391-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-392-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-393-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-395-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-397-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-400-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-403-331 + 53.00 104.43 1.1e+03
/cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-152-405-331 + 53.00 104.43 1.1e+03

seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:US-09-540-209B-3042
seq_documentation_block:
; Sequence 3042, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAGILIS
; FILE REFERENCE: 2709,1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3042
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: B. fragilis
; US-09-540-209B-3042

alignment_scores:
  Quality: 79.50      Length: 79
  Ratio: 1.728      Gaps: 2
Percent Similarity: 58.228      Percent Identity: 34.177

alignment_block:
US-09-528-682-4 x US-09-540-209B-3042 ..
Align seg 1/1 to: US-09-540-209B-3042 from: 1 to: 1656

12 GluProLeuAlaAsnThrAlaSerGlnPheLeuAsnMetGlyAs 28
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
97 CAGCGGGAAGGTAATCTCGCACATACATCTATTTTGAAT.....GA 140
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

28 nGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIleSerAla 45
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
141 TGAACAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 190
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

45 luGlyGlyMetAspArgAspLeuPhe.....GlySerAlaAsnIle 58
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
191 AAAAAGGCTTTGGAGCGTGCAGTGTGTTGGGAACAAGTGTGCTGTGAT 240
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

59 AspGlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysAl 75
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
241 ATAGTATGGCTAAATCACGCTGGTGTTCACATCGTTAGCTACTTAAC 290
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

75 aserSerAspThrAlaSerLeuAsnLysGlnHisasp 87
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
291 TAACGGTGAAGATGCCATCATGATGATGATGATGATGATGATGATGAT 327
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:US-09-540-209B-3340
seq_documentation_block:
; Sequence 3340, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAGILIS
; FILE REFERENCE: 2709,1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
```

```
; SEQ ID NO 3340
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: B. fragilis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2997),(3209)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are un
US-09-540-209B-3340

alignment_scores:
  Quality: 78.00      Length: 112
  Ratio: 1.322      Gaps: 7
Percent Similarity: 52.679      Percent Identity: 29.464

alignment_block:
US-09-528-682-4 x US-09-540-209B-3340 ..
Align seg 1/1 to: US-09-540-209B-3340 from: 1 to: 3381

10 AlaTyrGluProIleAlaAsnThrAsnThrThrGlnPheLeuAsn... 25
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2281 GCTTATCTGGCAATCGTGAATTTGACTGAAAGCCGCTTATATACACCGG 2330
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

26 ....MetGlyAsnGluValAlaLeuAspGly.....ArgTyrS 37
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2331 ATATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2380
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

37 erAsnTyrAlaLeuIleSerAlaGluGly.....GlyMet 48
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2381 AGAATAATGAACCTAGCTGGGAAGTGAATAAATAAATAAATAAATAA 2430
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

49 AspArgAspLeuPheGlySerAlaAsnIle.....AspGlyPheProG 63
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2431 GATTACAGCTCTCAATTTCAATTAATATACATGCTGACGGATTCAGG 2480
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

63 uValArg.....GluPheAsnSerLeuProAsnAsnLysA 75
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2481 GATTAGAGATAATATTTTTCAGCAGAGAAATTCATACCGAATATATCT 2530
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

75 laSerSerAspThrAlaSerLeuAsnLysGlnHisaspAlaAspPheLys 91
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2531 GAACCTCCAGTACTAAATCTAC.....GCCAATTCGCT 2565
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

92 LysTyrIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 103
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
2566 AAA.....GTGAAGAATAACAGGATTT 2586
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:US-09-540-209B-517
seq_documentation_block:
; Sequence 517, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAGILIS
; FILE REFERENCE: 2709,1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 517
; LENGTH: 1824
; TYPE: DNA
; ORGANISM: B. fragilis
; US-09-540-209B-517

alignment_scores:
  Quality: 72.50      Length: 114
  Ratio: 1.189      Gaps: 3
Percent Similarity: 53.509      Percent Identity: 23.684
```



[illegible]



```

; APPLICANT: Matsuda, Seichi P.T.
; APPLICANT: Hart, Elizabeth A.
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
; FILE REFERENCE: P02080US1/10025547
; CURRENT APPLICATION NUMBER: US/10/041,018
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259880
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-10-041-018-103

alignment_scores:
  Quality: 69.50      Length: 132
  Ratio: 1.053        Gaps: 8
  Percent Similarity: 50.000      Percent Identity: 27.273

alignment_block:
US-09-528-682-4 x US-10-041-018-103/rev ..

Align seg 1/1 to reverse of: US-10-041-018-103 from: 1 to: 2869

3  PheThrValArgIleGlnAspAlaTyrGluProIleAlaAsn.....Th 17
563  TTCACCAAAATGGGACAAAGCTTTAGATTCTATTATTATTCATGCGGA 514
17  rAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspG 34
513  TGAACATACACAGACGCTCTTACTCTAGTTACCAAGATGCAGTTGAT. 465
34  lyArgTyrSerAsnTyr..... 39
464  ....TATAATCGTTATGTCGGTAATATTACACAGGTCCTTATATTTA 420
40  AlaLeuIleSer.....AlaGluGlyGlyMetAspAr 50
419  AGTCATCATCTTTATTAGAACACGCTGATTTAAAGCGGACAAACGAT 370
50  gAspLeuPhe.....GlySerAlaAsn..... 57
369  TGGCTCTTTAGTTAGTTGTTCTGTTCTGTAGCGAGTCTTTAGTGGA 320
58  ....IleAspGlyPheProGlu.....ValArgGluPheAsnSer 69
319  CATTAGTAGATGGATTCAAGAGCAATTAGATGTTGAGCGCCACAAATCT 270
70  LeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHi 86
269  TTATTAATATAGAAATAGAGGTTCTGTT.....GATGAATA 232
86  sAspAlaAspPheLysLysTyrIleLysLeuLeuIleAsnAsnAsp 101
231  TGAACATTTCTTCAACAGCTTTGACCAATTAGAAATTGAATCATGAA 186

seq_name: /cqn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-8479

seq_documentation_block:
; Sequence 8479, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica

```

```

; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Lomelli, Michelle
; APPLICANT: Laroya, Mimi
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8479
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-544-8479

alignment_scores:
  Quality: 69.00      Length: 72
  Ratio: 1.683        Gaps: 4
  Percent Similarity: 56.944      Percent Identity: 31.944

alignment_block:
US-09-528-682-4 x US-09-721-544-8479 ..

Align seg 1/1 to: US-09-721-544-8479 from: 1 to: 365

13  ProIleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnG 29
75  CCGCGAGATTCACAAATCCAGTGACACATCTACACAATGATTAATCC 124
29  uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeu..... 41
125  AGTCCGCGCTGGAGGAGCGCGGTCCTCAACTTCCCGATGGTCCGCGCTCGG 174
42  ....IleSerAlaGluGlyGlyMetAsp.....ArgAsp 51
175  ACGGTCCGATGGCGCGATGGGTGGCATGGCCACACCATGAATGGA 224
52  LeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheAs 68
225  TCATTAGGGTCAGGCGACATAGACGACTTCCAAA.....AA 262
68  nSerLeuProAsnAsn 73
263  TTCT...CCTAACAC 275

seq_name: /cqn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-2707

seq_documentation_block:
; Sequence 2707, Application US/09919002
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin

```







```
13 ProIleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnG1 29
||| :|||||: : : : : : : : : : : : : : : : : : :
938 CCCGAGATTCAACAATTCAGTGACAACATCTACACATGATTATCC 987
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
29 uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeu. .... 41
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
988 AGTCCGCCTGGAGGACGCGGTCCAACTTCCCGATGGTCCCGGCTCGG 1037
42 .....IleSerAlaGluGlyMetAsp.....ArgAsp 51
||| :|||||: :|||||: :|||||: :|||||: :|||||:
1038 ACGTCCGATGGCGGCGATGGTGGCATGGAGCCACACCATGAATGGA 1087
52 LeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheAs 68
:|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
1088 TCATTAGGTCAGCGACATAGACGACTTCCAAA.....AA 1125
68 nSerLeuProAsnAsn 73
||| :|||||:
1126 TTCT...CCTAACAAAC 1138

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-500

seq_documentation_block:
; Sequence 500, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; PRIORITY FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 500
; LENGTH: 3303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2355)..(2355)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-105-299-500

alignment_scores:
Quality: 69.00 Length: 72
Ratio: 1.683 Gaps: 4
Percent Similarity: 56.944 Percent Identity: 31.944

alignment_block:
US-09-528-682-4 x US-10-105-299-500 ..
Align seg 1/1 to: US-10-105-299-500 from: 1 to: 3303

13 ProIleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnG1 29
||| :|||||: : : : : : : : : : : : : : : : : : :
938 CCCGAGATTCAACAATTCAGTGACAACATCTACACATGATTATCC 987
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
29 uValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeu. .... 41
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
988 AGTCCGCCTGGAGGACGCGGTCCAACTTCCCGATGGTCCCGGCTCGG 1037
42 .....IleSerAlaGluGlyMetAsp.....ArgAsp 51
||| :|||||: :|||||: :|||||: :|||||: :|||||:
1038 ACGTCCGATGGCGGCGATGGTGGCATGGAGCCACACCATGAATGGA 1087
52 LeuPheGlySerAlaAsnIleAspGlyPheProGluValArgGluPheAs 68
:|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
1088 TCATTAGGTCAGCGACATAGACGACTTCCAAA.....AA 1125
68 nSerLeuProAsnAsn 73
||| :|||||:
1126 TTCT...CCTAACAAAC 1138
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-935-625-13145

seq_documentation_block:
; Sequence 13145, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; PRIORITY FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 13145
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Misc.feature
; LOCATION: 1..4944
; OTHER INFORMATION: Ceres Seq. ID no. 3037797
US-09-935-625-13145

alignment_scores:
Quality: 69.00 Length: 90
Ratio: 1.211 Gaps: 2
Percent Similarity: 63.333 Percent Identity: 25.556

alignment_block:
US-09-528-682-4 x US-09-935-625-13145 ..
Align seg 1/1 to: US-09-935-625-13145 from: 1 to: 4944

21 ThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgTyrSe 37
||||| :|||||: :||| :||| :||| :||| :||| :||| :||| :|||
3184 ACGCAGAGACTCGACTGGGA.....GCTTGTCTGCTGTGTTTC 3224
37 rAsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG 54
||||| :|||||: :||| :||| :||| :||| :||| :||| :||| :|||
3225 AAAAGATATACAATTAGGAGGAGAAAGAGCTTTGATTCATTCGATTG 3274
54 lySerAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLeu 70
:|||||: :||| :||| :||| :||| :||| :||| :||| :||| :|||
3275 CTTCTACTAAACGCTGAAGCTGTCCCAAGATTTTACGAGATCTCTCTTG 3324
71 ProAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisAs 87
||||| :|||||: :||| :||| :||| :||| :||| :||| :||| :|||
3325 AGTAATCGTCTCTTTCTCCAGTGTGGCCTAAAGTATCAAGAGGA 3374
87 pAlaAspPheLysTyrIleLysLeuLeuIleAsnAspGlyPheP 104
:|||||: :||| :||| :||| :||| :||| :||| :||| :||| :|||
3375 AGACGACCTCGAA.....GACGGTTTCT 3397
104 heSerAsnAsnGlyGlyLys 110
: : : : : : : : : : : : : : : : : : : : : : : : :
3398 CAGAACTGGAAGGTTCAAAG 3417

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-25380

seq_documentation_block:
; Sequence 25380, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
```











```

38 AsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG1 54
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
604 AATATCTCCCTACACAGCTCATCTCCATGATGTCTCATCAGACCTCTCNT 555
54 ySerAlaAsnIleAsp.....GlyP 61
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
554 AGGATCCATCTAGTGGATCTTGACACATGCCTCTTAACATGTCCACTGGCT 505
61 heProGluValArgGluPheAsn..... 68
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
504 TTCCATTATCAGCAAAATAAAACCCATACACCTTATACAAAAGACCCCTAC 455
69 .....SerLeuProAsnAsnLysAlaSe 76
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
454 TTGATCTGGTCCAAATCTCTCAATTGCTCTCTCCACAACTGTATTTC 405
76 rSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysT 93
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
404 CAGTGATTTCCAGTCACATAACATGTAGCAT.....T 373
93 yTleLysLeuLeuIleAsnAsnAspGlyPhePhe 104
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
372 TTCTGAGTTTGCTCAATTTTAATAATATACATTTTC 338

```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-43500

```

seq_documentation_block:
; Sequence 43500, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43500
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(618)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-43500

```

```

alignment_scores:
  Quality: 66.00      Length: 95
  Ratio: 1.571      Gaps: 3
  Percent Similarity: 44.211      Percent Identity: 28.421

alignment_block:
US-09-528-682-4 x US-10-027-632-43500/rev ..

Align seg 1/1 to reverse of: US-10-027-632-43500 from: 1 to: 618

38 AsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG1 54
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
604 AATATCTCCCTACACAGCTCATCTCCATGATGTCTCATCAGACCTCTCNT 555
54 ySerAlaAsnIleAsp.....GlyP 61
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
554 AGGATCCATCTAGTGGATCTTGACACATGCCTCTTAACATGTCCACTGGCT 505
61 heProGluValArgGluPheAsn..... 68
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
504 TTCCATTATCAGCAAAATAAAACCCATACACCTTATACAAAAGACCCCTAC 455
69 .....SerLeuProAsnAsnLysAlaSe 76
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
454 TTGATCTGGTCCAAATCTCTCAATTGCTCTCTCCACAACTGTATTTC 405
76 rSerAspThrAlaSerLeuAsnLysGlnHisAspAlaAspPheLysLysT 93
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
404 CAGTGATTTCCAGTCACATAACATGTAGCAT.....T 373
93 yTleLysLeuLeuIleAsnAsnAspGlyPhePhe 104
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
372 TTCTGAGTTTGCTCAATTTTAATAATATACATTTTC 338

```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-43501

```

seq_documentation_block:
; Sequence 43501, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43501
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(618)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-43501

```

```

alignment_scores:
  Quality: 66.00      Length: 95
  Ratio: 1.571      Gaps: 3
  Percent Similarity: 44.211      Percent Identity: 28.421

alignment_block:
US-09-528-682-4 x US-10-027-632-43501/rev ..

Align seg 1/1 to reverse of: US-10-027-632-43501 from: 1 to: 618

38 AsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAspLeuPheG1 54
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```



















```
alignment_scores:
  Quality: 64.00      Length: 68
  Ratio: 1.641        Gaps: 4
  Percent Similarity: 57.353  Percent Identity: 32.353

alignment_block:
US-09-528-682-4 x US-09-539-331D-15353 ..
Align seg 1/1 to: US-09-539-331D-15353 from: 1 to: 267

17 ThrAsnThrThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAs 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 ACAAAATCCAGTGCACAAATATACAAATGATTAATCCAGTCCGCGCTGG 50
33 pGlyArgTyrSerAsnTyrAlaLeu.....IleS 43
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 AGGCAGCGCGTCCAACTCCCGATGGTCCGCGCTCNGACGGTCCGATGG 100
43 exAlaGluGlyGlyMetAsp.....ArgAspLeuPheGlySer 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 CGGCGATGGTGGCATGGAGCCACACACATCAATGGATCATTAGGGTCA 150
56 AlaAsnIleAspGlyPheProGluValArgGluPheAsnSerLeuProAs 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 GCGGACATAGCGGACTTCCAAA.....AATTCT...CCTAA 185
72 nAsn 73
|||||
186 CAAC 189
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-540-209B-3507

```
seq_documentation_block:
; Sequence 3507, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3507
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-3507
```

```
alignment_scores:
  Quality: 64.00      Length: 49
  Ratio: 1.882        Gaps: 0
  Percent Similarity: 69.386  Percent Identity: 30.612

alignment_block:
US-09-528-682-4 x US-09-540-209B-3507 ..
Align seg 1/1 to: US-09-540-209B-3507 from: 1 to: 1185

3 PheThrValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnTh 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
613 TGGACCAATCAAAATGAGCGCCATCTATTCCAGCCTGACCCGAAGCACT 662
19 rThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeuAspGlyArgT 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
663 GACCCACAGCCTATCTAAGCAACGACGAGGAGGAGCATCTATACCCGCA 712
36 yrSerAsnTyrAlaLeuIleSerAlaGluGlyMetAspArgAsp 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
713 TTTTCATCTACTCCCGCATACAAAGTGAAACGACGACGCGGGGAC 759
```

seq\_name: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-882-227-619

```
seq_documentation_block:
; Sequence 619, Application US/09882227
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 619
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)...(2274)
US-09-882-227-619
```

```
alignment_scores:
  Quality: 64.00      Length: 107
  Ratio: 1.255        Gaps: 6
  Percent Similarity: 47.664  Percent Identity: 24.299
```

alignment\_block:  
US-09-528-682-4 x US-09-882-227-619 ..

Align seg 1/1 to: US-09-882-227-619 from: 1 to: 2329

```
11 TyrGluProIleAla...AsnThrAsnThrThrThrGlnPhe.....Le 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
892 TATGAACCTTCTCCTTAAACACACCGATTTAATCCTTCGCTTATGT 941
24 uAsnMetGlyAsnGluValAlaLeu.....A 33
|||||:|||||:|||||:|||||:|||||:|||||:
942 GGATTTGGTAATGATGGTGGTTTGACCGAAGACATACTCTCAAAATG 991
33 spGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMetAsp 49
|||||:|||||:|||||:|||||:|||||:|||||:
992 ACACACGCCCTAAAGGCGATGGCTAGTGGAGGGGATTTTTC 1041
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgL 66
|||||:|||||:|||||:|||||:|||||:|||||:
1042 ACTCAATCTTTGGATTAGCAAGCTC...GTGTTCCCTGAA..... 1080
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83
|||||:|||||:|||||:|||||:|||||:|||||:
1081 .....AGACCTAATGAAAA..... 1095
83 snLysGlnHisAspAlaAspPheLysLysTyrIleLysLeuIleAsn 99
1095 ..... 1095
100 AsnAspGlyPhePheSerAsn 106
|||||:|||||:|||||:|||||:|||||:
1096 ...GATCCTTCTTTAGCAAT 1113
```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-027-632-110883

```
seq_documentation_block:
; Sequence 110883, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```



```

; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 110884
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-110884

```

alignment_scores:		
Quality:	64.00	Length: 102
Ratio:	1.164	Gaps: 3
Percent Similarity:	53.922	Percent Identity: 23.529

US-09-528-682-4 x US-10-027-632-110884

Align seq 1/1 to: US-10-027-632-110884 from: 1 to: 2354

16 AsnThrAsnThrThrGlnPheLeuAsnMetGlyAsnGluValAlaLeu 32  
||||| :||| ||| :||| :|||  
780 ACACAAATAAAAAATACTCCTTTCTCTAATATTGCGAAAANAATAGACCTT 829

32 uAspGlyArgTyrSerAsnTyrAlaLeuIleSerAlaGluGlyGlyMeta 49

| |||||:: ::::: :::: :

830 GTACAGAACCTATTCAACTACAGTAACAACCTCTCATCATGAACAACGGA 879

49 spArgAspLeupheGlySerAlaAsnIleAspGlyPheProGluValArg 65  
 :||:|||||  
 880 ATAGAAATATGTTT.....TCTGGATATCCACACATCAGA 914

[illegible]

76 .....SerSerAspThrAlaSerLeuAsnLysGlnH 86

959 ATAGCATTCCTCTGTAGTTTGAAACTATTTGTTCTCTCAATGACTTCTGCAC 100

[illegible]

103 PhePhe 104

1059 T T C T T T 1064

seq\_name: /cqn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:US-09-975-254-29497

```
seq_documentation_block:
; Sequence 29457, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molec
```

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome



```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 29497
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700900007H1
US-09-975-254-29497
```

```
alignment_scores:
  Quality: 63.50      Length: 73
  Ratio: 1.587       Gaps: 1
Percent Similarity: 54.795 Percent Identity: 28.767

alignment_block:
US-09-528-682-4 x US-09-975-254-29497 ..
Align seg 1/1 to: US-09-975-254-29497 from: 1 to: 263

13 ProfileAlaAsnThrAsnThrThr.....GlnPheLeuAsnMe 26
   ||| |||::|||::: ||| |||::: |||::: |||
28 CCACAGGCTGATCTAGACTGTACACTTTCAGAGTCAGTTGAAGAGCAT 77

26 tGlyAsnGluValAlaLeuAspGlyArgTyrSerAsnTyrAlaLeuIles 43
   |||::: |||::: ||| |||::: ||| |||::: |||
78 GGGTCGTGGAGTTAGCAGTGTGGTGGTCAAGTTCCTTGGTTATCTTT 127

43 erAlaGluGlyGlyMetAspArgAspLeuPheGlySerAlaAsnIleasp 59
   ::::: ||| ::::: |||
128 TTGGCAGTGGAGAGCCCTGCCGCAAAACCGGCACAAACAATGCCCA 177

60 GlyPheProGluValArgGluPheAsnSerLeuProAsnAsnLysAlase 76
   :: ||| |||::: |||::: ||| |||::: |||
178 GCCCAACCTGAAGTCAGCGTGTAAATAATAGCGCTCTCTCAAAAGCATC 227

76 rSerAspThrAlaSerLeu 82
   ||| ||| :::::
228 TCCTCTACTAAACAATA 246
```

```
seq_name: /cgn2_6/ptodata/1/pna/us10_NEW_COMB.seq:US-10-027-632-268304

seq_documentation_block:
; Sequence 268304, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 268304
```

```
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-268304

alignment_scores:
  Quality: 63.50      Length: 67
  Ratio: 1.671       Gaps: 3
Percent Similarity: 56.716 Percent Identity: 28.358

alignment_block:
US-09-528-682-4 x US-10-027-632-268304 ..
Align seg 1/1 to: US-10-027-632-268304 from: 1 to: 523

36 TyrSerAsnTyrAlaLeuIlesSerAlaGluGlyGlyMetAspArgAspLe 52
   |||::: ||| |||::: |||::: |||::: |||
297 TATCCAGATTCTGCCCTGGAGGCGCTGCAGGCGTGTGTTCAGGAAGCT 346

52 uPhe.....GlySerAlaAsn...IleAspGlyP 61
   ||| ::::: |||
347 GTTTGAAGATCAGACGAGAGCCCTGTCCACAGCAAACTATAAATGATGCT 396

61 heProGluValArgGluPheAsnSerLeuProAsnAsnLysAlaSerSer 77
   :: ::::: ||| |||::: |||::: |||
397 ATGTGGACATGAGCAGCTTCAACACCTTTAAGAACAATCAGCAGAGCACA 446

78 AspThrAlaSerLeuAsnLysGlnHisAlaAspPheLysLysTyrIle 94
   ||| ||| ::::: |||
447 .....GACCAGGACGCAAGAGGTACCT 469

94 e 94
;
470 G 470
```

```
seq_name: /cgn2_6/ptodata/1/pna/us09_NEW_COMB.seq:US-09-620-393B-4281

seq_documentation_block:
; Sequence 4281, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4281
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..577
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..577
; OTHER INFORMATION: Ceres Seq. ID 1387475
US-09-620-393B-4281
```

```
alignment_scores:
  Quality: 63.50      Length: 77
  Ratio: 1.984       Gaps: 4
Percent Similarity: 41.558 Percent Identity: 28.571

alignment_block:
US-09-528-682-4 x US-09-620-393B-4281 ..
Align seg 1/1 to: US-09-620-393B-4281 from: 1 to: 577
```











Percent Similarity: 57.798 Percent Identity: 25.688

alignment\_block:

US-09-528-682-4 x US-10-128-714-2019/rev ..

Align seg 1/1 to reverse of: US-10-128-714-2019 from: 1 to: 1263

```
5 ValArgileGlnAspAlaTyrGluProIleAlaAsnThrAsnThrThr 21
||||:||||| :||| |||:| |||:|||||
726 GTTCAGCTTCAGCACACCATGAGGATAGAGCCACCTGCACACCCCT 677
21 rGlnPheLeuAsnMetGlyAsnGlu...ValAlaLeuAspGlyArgTyrS 37
||||:||||| :||| |||:| |||:|||||
676 TCAACTCATCGATCTCGCACCGGTTTGTGACGTCGAAGGAACGAATG 627
37 erAsnTyrAlaLeuIleSerAlaGlu.....GlyGlyMetAsp... 49
||||:||||| :||| |||:| |||:|||||
626 ACAATCATGTGAGCGGAAGCAGTGAAGTCACGGACGGAATGGGATGTG 577
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG 66
||||| :||| :||| :||| :|||
576 CGAGAC.....AAGGTATTCTGTGACAGCGTCGATGTTGACTTGA 536
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeu 83
||||:||||| :||| |||:| |||:|||||
535 GCTGACGACAGATGGGAATGATAGGAGAGCCATCGGCAACAGTACCAC 486
83 snLysGlnHisAsp.....AlaAspPheLysLys 92
||||:||||| :||| |||:| |||:|||||
485 ATGAACCTTCAGGATTGATGTTGTAATGTTGCAGAGCACCGCTTCCCT 436
93 TyrIleLysLeuLeuIleAsnAsnAsp 101
||||:||||| :||| |||:| |||:|||||
435 CAGATCAACCTTGTTCGACAGATGAT 409
```

seq\_name: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:PCT-US02-13142-1019

seq\_documentation\_block:

```
; Sequence 1019, Application PC/TUS0213142
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-228
; CURRENT APPLICATION NUMBER: PCT/US02/13142
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1019
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
PCT-US02-13142-1019
```

alignment\_scores:

Quality: 63.50 Length: 109  
Ratio: 1.008 Gaps: 5  
Percent Similarity: 57.798 Percent Identity: 25.688

alignment\_block:

US-09-528-682-4 x PCT-US02-13142-1019/rev ..

Align seg 1/1 to reverse of: PCT-US02-13142-1019 from: 1 to: 1542

```
5 ValArgileGlnAspAlaTyrGluProIleAlaAsnThrAsnThrThr 21
||||:||||| :||| |||:| |||:|||||
948 GTTCAGCTTCAGCACACCATGAGGATAGAGCCACCTGCACACCCCT 899
21 rGlnPheLeuAsnMetGlyAsnGlu...ValAlaLeuAspGlyArgTyrS 37
||||:||||| :||| |||:| |||:|||||
898 TCAACTCATCGATCTCGCACCGGTTTGTGACGTCGAAGGAACGAATG 849
37 erAsnTyrAlaLeuIleSerAlaGlu.....GlyGlyMetAsp... 49
||||:||||| :||| |||:| |||:|||||
848 ACAATCATGTGAGCGGAAGCAGTGAAGTCACGGACGGAATGGGATGTG 799
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG 66
||||| :||| :||| :||| :|||
798 CGAGAC.....AAGGTATTCTGTGACAGCGTCGATGTTGACTTGA 758
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeu 83
||||:||||| :||| |||:| |||:|||||
757 GCTGACGACAGATGGGAATGATAGGAGAGCCATCGGCAACAGTACCAC 708
83 snLysGlnHisAsp.....AlaAspPheLysLys 92
||||:||||| :||| |||:| |||:|||||
707 ATGAACCTTCAGGATTGATGTTGTAATGTTGCAGAGCACCGCTTCCCT 658
93 TyrIleLysLeuLeuIleAsnAsnAsp 101
||||:||||| :||| |||:| |||:|||||
657 CAGATCAACCTTGTTCGACAGATGAT 631
```

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-1019

seq\_documentation\_block:

```
; Sequence 1019, Application US/10128714
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1019
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-1019
```

alignment\_scores:

Quality: 63.50 Length: 109  
Ratio: 1.008 Gaps: 5  
Percent Similarity: 57.798 Percent Identity: 25.688

alignment\_block:

US-09-528-682-4 x US-10-128-714-1019/rev ..



5 ValArgIleGlnAspAlaTyGluProIleAlaAsnThrAsnThrThr 21  
||||:||||| :: ||| |||:: |||||  
1014 GTTCAGGTTGACGACACGAGTGAGGATAGAGCCACCTGCGACACACCCCT 965







```

37 eRAsnTyrAlaLeuIleSerAlaGlu.....GlyGlyMetAsp... 49
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
980 ACAATCATCTGAGGGCAACAGCAGTGAACTCACGCAGCGGAACCTGGGTG 931
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
50 AtgAspLeuPheGlySerAlaAlanIleAspGlyPheProGluValArgGl 66
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
930 CGAGAC.....AAGGTATTGCTGTGACAGCGTCGATGTTGTACTTGA 890
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
889 GCTGACGAGATGGCAATGATAGGAGACCATCGCCACAGTACCACCA 840
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
83 snLysGlnHisasp.....AlaAspPheLysLys 92
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
839 ATGAACCTTCAGGATTGATTGTTAAATGTCGAGAGCACCGCTCTCCCTCAT 790
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
93 TyrIleLysLeuLeuIleAsnAsnasp 101
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
789 CAGATCAACCTTGTCTTCGAGAATGAT 763
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq;US-10-106-698-225

seq_documentation_block:
; Sequence 225, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleo
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIORITY FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 225
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-225

alignment_scores:
Quality: 63.50 Length: 114
Ratio: 1.058 Gaps: 5
Percent Similarity: 52.632 Percent Identity: 21.930

alignment_block:
US-09-528-682-4 x US-10-106-698-225 ..

Align seg 1/1 to: US-10-106-698-225 from: 1 to: 1762

11 TyrGluProIleAlaAsnThrAsnThrThrGlnPheLeuAsnMetGl 27
|||||::|||::|||::|||::|||::|||::|||::|||::|||
744 TACCAGCCATCCCGGCGGG.....CTCAACGTGGG 775

27 yAnsgluValAlaLeuAspGlyArgTyrSerAsnTyrAla..... 40
|||||::|||::|||::|||::|||::|||::|||::|||::|||
776 AATGTCTGTTCATCCCAAGGAGTGTCTCAGCAGCACATGAAGCGGTTCT 825

41 .....LeulleSerAlaGluGlyGlyMetAspArgaspleuPhe 53
|||||::|||::|||::|||::|||::|||::|||::|||::|||
826 TCGTGAACATTTCTGTGTGGGAGGATCGCGGCTCAGACATCGCCTTCCAC 875

54 GlySerAlaAsnIleAspGlyPheProGluValArgGluPheAsnSerIe 70
|||||::|||::|||::|||::|||::|||::|||::|||::|||
876 TTCATCCGGGTTTGACGGCTGGGACACAGGTG...GTCTTCAACACGTT 922

70 uProAsnLysAlaSerSerAspThrAlaSerLeuAsnLysGlnHisa 87
|||||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

923 GCAGGCGGGGAAGTGGGGCAGCAG.....GAGAGGAAGAGGA 960
      87 spAlaAspPheLys.....
      :  ||| ||| |||
961 GCATGCCCTCAAAAGGTTGCCCTTTGAGCTGGTCTTCATAGTCTGT 1010
      :  ||| ||| ||| ||| ||| ||| ||| |||
1011 GCTGAGCACTACAAGGTGGTGAATGGAATCCCTTCTAT 1052

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq: PCT-US02-13142-19
seq_documentation_block:
; Sequence 19, Application PC/TUS0213142
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Asperg
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-228
; CURRENT APPLICATION NUMBER: PCT/US02/13142
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 3542
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
PCT-US02-13142-19

alignment_scores:
  Quality: 63.50      Length: 109
  Ratio: 1.008        Gaps: 5
  Percent Similarity: 57.798      Percent Identity: 25.688

alignment_block:
US-09-528-682-4 x PCT-US02-13142-19/rev ..

Align seg 1/1 to reverse of: PCT-US02-13142-19 from: 1 to: 3542

5 ValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnThrTh 21
|||||:||||| :||| |||:|||||
1948 GTTCACCTTCAGCACCATGAGTAGGATGAGCCACCTGCAACACCCCT 1899
|||||:||||| :||| |||:|||||
21 rGlnPheLeuAsnMetGlyAsnGlu...ValAlaLeuAspGlyArgTyr 37
|||||:||||| :||| |||:|||||
1898 TCAACTCATCGATCTCCGACCGGGTTGTGTGACGTGCAAGAACGATG 1849
|||||:||||| :||| |||:|||||
37 exAsnTyrAlaLeuIleSerAlaGlu.....GlyGlyMetasp... 49
|||||:||||| :||| |||:|||||
1848 ACAATCATGTAGCGCAAGCAGTGAAGTCACGGACGGAACTGGGATGTG 1799
|||||:||||| :||| |||:|||||
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArg 66
|||||:||||| :||| |||:|||||
1798 CGAGAC.....AAGGTATTCTGTGACAGCGTCGATGTTGTACTTGA 1758
|||||:||||| :||| |||:|||||
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeu 83
|||||:||||| :||| |||:|||||
1757 GCTGACGACAGATGGGAATGATAGGAGAGCCATCGGCACAGTACCAGA 1708
|||||:||||| :||| |||:|||||
83 snLysGlnHisasp.....AlaAspPheLysLys 92
|||||:||||| :||| |||:|||||
1707 ATGCAACTTCAGGATTTGTTGGTAAATGTTGCAGACGACCGTCTTCCTCAT 1658
|||||:||||| :||| |||:|||||

```



93 TyrIleLysLeuLeuIleAsnAsp 101  
|||:|||||:|||||  
1657 CAGATCAACCTTGTTCGAGATGAT 1631

93 TyrIleLysLeuLeuIleAsnAsp 101  
|||:|||||:|||||  
1657 CAGATCAACCTTGTTCGAGATGAT 1631

seq\_name: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:US-10-128-714-19

seq\_documentation\_block:  
; Sequence 19, Application US/10128714  
; GENERAL INFORMATION:  
; APPLICANT: Jlang, Bo  
; APPLICANT: Hu, Wengqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lenieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 3542  
; TYPE: DNA  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-19

alignment\_scores:  
Quality: 63.50 Length: 109  
Ratio: 1.008 Gaps: 5  
Percent Similarity: 57.798 Percent Identity: 25.688  
alignment\_block:  
US-09-528-682-4 x US-10-128-714-19/rev ..  
Align seg 1/1 to reverse of: US-10-128-714-19 from: 1 to: 3542  
5 ValArgIleGlnAspAlaTyrGluProIleAlaAsnThrAsnThrThrTh 21  
|||:|||||:|||||  
1948 GTTCAGCTTCAGCACACCATGAGATAGAGCCACCTGCAACACCCCT 1899  
|||:|||||:|||||  
21 rGlnPheLeuAsnMetGlyAsnGlu...ValAlaLeuAspGlyArgTyrS 37  
|||:|||||:|||||  
1898 TCAACTCATCGATCTCCGACCGGGTTTGTGACGTCGAAGGACGATG 1849  
|||:|||||:|||||  
37 erAsnTyrAlaLeuIleSerAlaGlu.....GlyGlyMetAsp... 49  
:|||||:|||||  
1848 ACAATCATGTGAGCGGAGCAGTGAAGTCACGGACGGGAACTGGGATGTG 1799  
|||:|||||:|||||  
50 ArgAspLeuPheGlySerAlaAsnIleAspGlyPheProGluValArgG 66  
|||||:|||||:|||||  
1798 CGAGAC.....AAGGTATTCTGTTGACAGCGTCGATGTTGTTACTGA 1758  
|||:|||||:|||||  
66 uPheAsnSerLeuProAsnAsnLysAlaSerSerAspThrAlaSerLeuA 83  
|||:|||||:|||||  
1757 GCTGACGACAGATGGGATGATAGGAGAGCCATCGGCAACAGTACCACGA 1708  
|||:|||||:|||||  
83 snLysGlnHisAsp.....AlaAspPheLysLys 92  
|||:|||||:|||||  
1707 ATGAACCTCAGGATTGATTGGTAAATGTTGCAGACGACCGTCTTCCTCAT 1658  
|||:|||||:|||||



